

PA (STER-) STERRENBEID BIOTECHNOLOGIE NORTH AMERICA.
 XX Carcagno CM, Criscuolo M, Melo C, Vidal JA;
 PI WPI: 2000-376485/32.
 XX

PT Novel methods for purifying recombinant human erythropoietin from
 XX mammalian cell culture reagents -
 XX Claim 16; Page 18; 30pp; English.

CC The present invention relates to a method for purifying erythropoietin
 CC (EPO) for treatment of disease, especially anaemia. The method involves
 CC treating cell culture supernatants with differential precipitation,
 CC hydrophobic interaction chromatography, diafiltration, anionic and
 CC cationic exchange chromatography and molecular exclusion
 CC chromatography. The present sequence is the protein from the culture
 CC supernatant of transfected cell lines, after purification by the above
 CC process. The sequence shows total homology with natural human EPO.
 CC The advantage of this method is that high purity and quality EPO is
 CC produced. A further advantage is that the process does not involve the
 CC use of organic solvents that may harm the environment.

SO Sequence 165 AA;

Query Match Best Local Similarity 100.0%; Score 846; DB 21; Length 165;
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVNFFYAMKRMVEVGOQA 60
 DB 1 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVNFFYAMKRMVEVGOQA 60
 OY 61 VEWMOGIALISEAVLRGQALLVNSQWPWEFLQHVDAVSGLSLTLRLALGAOKEAIS 120
 DB 61 VEWMOGIALISEAVLRGQALLVNSQWPWEFLQHVDAVSGLSLTLRLALGAOKEAIS 120
 OY 121 PPDASAAPIRTITADTFRKLFRRYSNPLRGKLTLYTGACRTGD 165
 DB 121 PPDASAAPIRTITADTFRKLFRRYSNPLRGKLTLYTGACRTGD 165

RESULT 2

ID AAB03760 standard; protein; 165 AA.

AC AAB03760;

DT 04-OCT-2000 (first entry)

DE Human erythropoietin (EPO) amino acid sequence.

KW Erythropoietin; EPO; human; erythroblast differentiation; anaemia;

OS Homo sapiens.

PN WO200027997-A1.

PD 18-MAY-2000.

PF 08-NOV-1999; 99WO-US26240.

PR 06-NOV-1998; 98AR-0105611.

PR 23-FEB-1999; 99AR-0100683.

PA (STER-) STERRENBEID BIOTECHNOLOGIE NORTH AMERICA.

PI Carcagno CM, Criscuolo M, Melo C, Vidal JA;

DR WPI: 2000-376519/32.

PT A novel method for the massive culture of recombinant mammalian cells

PT producing recombinant human erythropoietin -
 XX Example 8; Page 11-12; 23pp; English.

CC This sequence represents the human erythropoietin amino acid sequence.
 CC Erythropoietin is a glycoprotein that stimulates erythroblast
 CC differentiation in the bone marrow. The present invention relates to a
 CC method for the large scale production of human EPO from recombinant
 CC mammalian cells. The method comprises culturing mammalian cells which
 CC express recombinant human EPO in culture medium comprising insulin.
 CC Erythropoietin can be used to treat anaemia derived from renal failure.
 CC The method allows for the industrial scale production of EPO, and
 CC overcomes the problems of low reproducibility and output quality which
 CC are encountered with previous production methods.

SO Sequence 165 AA;

Query Match Best Local Similarity 100.0%; Score 846; DB 21; Length 165;
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVNFFYAMKRMVEVGOQA 60
 DB 1 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVNFFYAMKRMVEVGOQA 60
 OY 61 VEWMOGIALISEAVLRGQALLVNSQWPWEFLQHVDAVSGLSLTLRLALGAOKEAIS 120
 DB 61 VEWMOGIALISEAVLRGQALLVNSQWPWEFLQHVDAVSGLSLTLRLALGAOKEAIS 120
 OY 121 PPDASAAPIRTITADTFRKLFRRYSNPLRGKLTLYTGACRTGD 165
 DB 121 PPDASAAPIRTITADTFRKLFRRYSNPLRGKLTLYTGACRTGD 165

RESULT 3

ID AAY9705 standard; protein; 165 AA.

AC AAY9705;

DT 15-SEP-2000 (first entry)

DE Non-glycosylated erythropoietin analogue NGE-166delta.

KW Human; non-glycosylated erythropoietin analogue; NGEA; haematocrit;

OS antianemic; anaemia; erythropoiesis promoter; mutant; mutein.

OS Homo sapiens.

PN WO200032772-A2.

PD 08-JUN-2000.

PF 23-NOV-1999; 99WO-US27801.

PR 30-NOV-1998; 98US-0110289.

PA (ELIL) LILLY & CO ELI.

PI Beals JM, Glaesner W, Micranovic R, Milligan RL, Wlitcher DR;

DR WPI: 2000-412320/35.

DR N-PSDB; AAA48373.

PT Non-glycosylated erythropoietic compound useful for increasing

hematocrit level in mammal with insufficient hematocrit levels in
 conditions such as anemia, comprises protein covalently bonded to
 polymer -
 Claim 2; Page 93-94; 94pp; English.
 The present sequence is a non-glycosylated erythropoietin analogue

CC (NGBA) designated NGE-166delta. The protein sequence is identical to
CC the sequence of wild-type human non-glycosylated erythropoietin NGE
CC except that Arg at position 166 is deleted. NGE promotes erythropoiesis
CC and can therefore be used to increase haematocrit levels in mammals
CC with conditions such as anaemia, in which levels of haematocrit are
CC insufficient. NGE analogues can also be used to treat such conditions.
CC NGEAs do not themselves cause a significant increase in haematocrit but
CC they acquire that property once they are derivatised with polyethylene
CC glycol polymers. The analogues can be produced using a linkerless
CC aldehyde modification process. They show stability and bioactivity in
CC vivo. The nucleotide sequence encoding this protein was constructed
CC synthetically by in vitro hybridisation using a set of six overlapping
CC oligonucleotides from the positive strand of human erythropoietin cDNA
CC with six complementary oligonucleotides (negative strand). The codon
CC usage was 100% optimised for E. coli codon usage. The hybridised
CC oligonucleotides were ligated with T4 DNA ligase and the ligation product
CC amplified by PCR. The nucleotide sequence was used to express the protein
CC in host cells.

SQ Sequence 165 AA:

Query Match 100.0%; Score 846; DB 21; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLCDISRVLERYLLEKAEENITTCGAEHCSTLNEITVPTKYNFYAKRMKEVGQA 60
DB 1 APPRLCDISRVLERYLLEKAEENITTCGAEHCSTLNEITVPTKYNFYAKRMKEVGQA 60
QY 61 VEWMOGLALISEAVLNGQALLVNSSQPWEPLQLHVDKAVSGLSLTTLRALGAQKEAIS 120
DB 61 VEWMOGLALISEAVLNGQALLVNSSQPWEPLQLHVDKAVSGLSLTTLRALGAQKEAIS 120
QY 121 PPDASAAPLRTITADTFERKLFRRVYSNPLRGKIKLYTGACRTGD 165
DB 121 PPDASAAPLRTITADTFERKLFRRVYSNPLRGKIKLYTGACRTGD 165

RESULT 4
AAV93445

ID AAV93445 standard; protein; 165 AA.

XX AAV93445;

DT 04-SEP-2000 (first entry)

XX Amino acid sequence of human erythropoietin.

KW Human; erythropoietin; EPO; anaemia; renal failure.

OS Homo sapiens.

PN WO200028066-A1.

PD 18-MAY-2000.

PF 08-NOV-1999; 99WO-US26238.

PR 06-NOV-1998; 98AR-0105609.

PR 23-FEB-1999; 99AR-0100679.

PA (STER-) STERRENBEID BIOTECHNOLOGIE NORTH AMERICA.

PI Carcagno CM, Criscuolo M, Melo C, Vidal JA;

DR WPI: 2000-376574/32.

PT New host cell producing recombinant human erythropoietin (EPO) used for
large scale production of EPO -

PS Claim 1; Page 26-27; 51pp; English.

XX The present sequence represents human erythropoietin protein. The

CC specification describes a host cell line which is used to produce
CC human erythropoietin (EPO). EPO is a glycoprotein. The cell line
CC is used for the production of recombinant human erythropoietin. The
CC protein is used for the treatment of anaemia, especially anaemia derived
CC from renal failure.

SQ Sequence 165 AA:

Query Match 100.0%; Score 846; DB 21; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLCDISRVLERYLLEKAEENITTCGAEHCSTLNEITVPTKYNFYAKRMKEVGQA 60
DB 1 APPRLCDISRVLERYLLEKAEENITTCGAEHCSTLNEITVPTKYNFYAKRMKEVGQA 60
QY 61 VEWMOGLALISEAVLNGQALLVNSSQPWEPLQLHVDKAVSGLSLTTLRALGAQKEAIS 120
DB 61 VEWMOGLALISEAVLNGQALLVNSSQPWEPLQLHVDKAVSGLSLTTLRALGAQKEAIS 120
QY 121 PPDASAAPLRTITADTFERKLFRRVYSNPLRGKIKLYTGACRTGD 165
DB 121 PPDASAAPLRTITADTFERKLFRRVYSNPLRGKIKLYTGACRTGD 165

RESULT 5
AAB84525

ID AAB84525 standard; protein; 165 AA.

XX AAB84525;

DT 05-SEP-2001 (first entry)

XX Amino acid sequence of human erythropoietin (EPO) protein.

KW Erythropoietin; EPO; erythropoietin stimulating protein; NESP;

KW sustained release.

OS Homo sapiens.

PN WO200130320-A1.

PD 03-MAY-2001.

PF 23-OCT-2000; 2000WO-US29257.

PR 22-OCT-1999; 99US-0426566.

PR 13-OCT-2000; 2000US-2222222.

PA (AMGE-) AMGEN INC.

PI Burke P, Klumb L, Murphy K, Herberger J, French DL;

DR WPI: 2001-417552/44.

PT Sustained release composition comprises an active biological
ingredient, notably a protein or other biopolymer, particularly
erythropoietin stimulating protein, in biocompatible, biodegradable
polymeric microparticles -

PS Disclosure; Page 56; 61pp; English.

CC The present sequence encodes a human erythropoietin (EPO) protein.
CC The specification describes a composition for the sustained release of
CC biologically active EPO stimulating protein (NESP). The reduced
CC frequency of administration of NESP, which requires preferably injection
CC by skilled personnel, improves patient compliance. Also, sustained
CC release reduces the nature and severity of any side effects of the drug.

SQ Sequence 165 AA:

Query Match 100.0%; Score 846; DB 22; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;

	Matches	165,	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	APPRICISRVLYRTLEKAEKENTTTGCAEHCISNENTTADPTVYNFPAKRMREYGOQA	60							
Db	1	APPRICISRVLYRTLEKAEKENTTTGCAEHCISNENTTADPTKNEFPAKRMREYGOQA	60							
QY	61	VEWMOGLALISAVYLRGOALLVNSSQPMWPELOLHVDKAVSGIRSLFTTLRALGAKREKIS	120							
Db	61	VEWMOGLALISAVYLRGOALLVNSSQPMWPELOLHVDKAAVSGIRSLFTTLRALGAKREKIS	120							
QY	121	PPDAASAPLRTITADTEFKLEFRVYSNPLRGKILKLYTGACRGTG	165							
Db	121	PPDAASAPLRTITADTEFKLEFRVYSNPLRGKILKLYTGACRGTG	165							

XX	RESULT 6
XX	AAB66697
ID	AAB66697 standard; protein; 165 AA.
XX	
XX	AAB66697;
XX	
DT	06-APR-2001 (first entry)
XX	
DE	Human erythropoietin protein #1.
XX	
KW	Erythropoietin; EPO; reticulocytes; red blood cell;
KW	ethylene glycol; chronic renal failure; AIDS; cancer.
XX	
OS	Homo sapiens.
XX	
PX	WO200102017-A2.
PX	
PX	11-JAN-2001.
PX	
PF	28-JUN-2000; 2000WO-EP06009.
PX	
PA	02-JUL-1999; 99US-0142243.
PR	05-AUG-1999; 99US-0147452.
PR	30-AUG-1999; 99US-0151454.
XX	
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.
XX	
FI	Burg J, Hilger B, Josel H;
XX	
DR	WPI: 2001-147051/15.
XX	
FT	Novel erythropoietin-glycoprotein conjugate useful for treating
FT	diseases correlated with anemia in chronic renal failure patients, AIDS
FT	linker -
XX	
PS	Claim 19; Fig 1; 40pp; English.
XX	
CC	The present invention relates to a conjugate comprising, human
CC	erythropoietin glycoprotein (EPO) having at least one free
CC	amino group and having in vivo biological activity of causing
CC	an increase the production of reticulocytes and red blood cells,
CC	covalently linked to 1-3 lower-alkoxy poly(ethylene glycol) groups
CC	through a linker. The invention is useful for preparation of
CC	medicaments for the treatment of prophylaxis of disease correlated
CC	with anemia in chronic renal failure patients (CRF), AIDS and for
CC	the treatment of cancer patients undergoing chemotherapy.
XX	
XX	Sequence 165 AA;

QY	QY	QY	QY
61	61	61	61
VEWMGALLISFAYIRGALLVNSSQPHEPIQLHVADKVASLSRLTTLLEALGAKKAIS	VEWMGALLISFAYIRGALLVNSSQPHEPIQLHVADKVASLSRLTTLLEALGAKKAIS	VEWMGALLISFAYIRGALLVNSSQPHEPIQLHVADKVASLSRLTTLLEALGAKKAIS	VEWMGALLISFAYIRGALLVNSSQPHEPIQLHVADKVASLSRLTTLLEALGAKKAIS
120	120	120	120
Db	61	61	61
VEWMGALLISFAYIRGALLVNSSQPHEPIQLHVADKVASLSRLTTLLEALGAKKAIS	VEWMGALLISFAYIRGALLVNSSQPHEPIQLHVADKVASLSRLTTLLEALGAKKAIS	VEWMGALLISFAYIRGALLVNSSQPHEPIQLHVADKVASLSRLTTLLEALGAKKAIS	VEWMGALLISFAYIRGALLVNSSQPHEPIQLHVADKVASLSRLTTLLEALGAKKAIS
120	120	120	120
QY	121	121	121
PPDAASAPLRTITADPFRKLEFRVYSNFRGKLELYTGEACRTGD	PPDAASAPLRTITADPFRKLEFRVYSNFRGKLELYTGEACRTGD	PPDAASAPLRTITADPFRKLEFRVYSNFRGKLELYTGEACRTGD	PPDAASAPLRTITADPFRKLEFRVYSNFRGKLELYTGEACRTGD
165	165	165	165
Db	121	121	121
PPDAASAPLRTITADPFRKLEFRVYSNFRGKLELYTGEACRTGD	PPDAASAPLRTITADPFRKLEFRVYSNFRGKLELYTGEACRTGD	PPDAASAPLRTITADPFRKLEFRVYSNFRGKLELYTGEACRTGD	PPDAASAPLRTITADPFRKLEFRVYSNFRGKLELYTGEACRTGD
165	165	165	165

XX	RESULT 7
XX	ABB77896
XX	ABB77896 standard; protein; 165 AA.
XX	ABB77896
XX	ABB77896;
XX	07-OCT-2002 (first entry)
XX	
XX	Amino acid sequence of a human erythropoietin (Epo).
XX	
XX	Human; erythropoietin; EPO; glycoprotein; reticulocyte production;
XX	red blood cell production; anaemia; chronic renal failure;
XX	acquired immunodeficiency syndrome; AIDS; cancer; bone marrow;
XX	committed erythroid progenitor.
XX	
XX	Homo sapiens.
XX	
XX	WO200249673-A2.
XX	
XX	27-JUN-2002.
XX	
XX	08-DEC-2001; 2001WO-EP14434.
XX	
XX	20-DEC-2000; 2000EP-0127891.
XX	
XX	(HOFF) HOFFMANN LA ROCHE & CO AG F.
XX	
XX	Burg J, Engel A, Franze R, Hilger B, Schurig HE, Tischer W;
XX	Wozny M;
XX	
XX	WPI; 2002-566640/60.
XX	
XX	
XX	Claim 26; Fig 1; 40pp; English.
XX	
XX	The present sequence represents a human erythropoietin (Epo) protein.
XX	It was used to produce conjugates of the invention. The specification
XX	describes a conjugate comprising an EPO glycoprotein having an N-terminal
XX	alpha-amino group, chosen from human EPO (hEPO) or its analogues (where
XX	hEPO is modified by addition of 1-6 glycosylation sites or a
XX	rearrangement of a glycosylation site). The glycoprotein is covalently
XX	linked to a poly(ethylene glycol) group. The EPO glycoprotein has in vivo
XX	biological activity of causing bone marrow cells to increase production
XX	of reticulocytes and red blood cells. The conjugate increased circulating
XX	half-life and plasma residence time, decreased clearance, increased
XX	clinical activity in vivo, improved potency and stability, when compared
XX	to unmodified EPO. The EPO conjugate is useful for preparing medicaments
XX	for the treatment and prophylaxis of diseases correlated with anaemia in
XX	chronic renal failure patients (CRF), acquired immunodeficiency syndrome
XX	(AIDS) and for treating cancer patients undergoing chemotherapy. It is
XX	also useful for treating patients by stimulating the division and
XX	differentiation of committed erythroid progenitors in the bone marrow.
XX	
XX	Sequence 165 AA:
XX	

Query Match	Similarity	Score	846;	DB	22;	Length	165;
Best Local	Similarity	100.0%;	Pred.	No.	1.1e-86;		
Matches	165;	Conservative	0;	Mismatches	0;	Indels	0;
1	APPLICDSRYLEERYLLLEAKAEENITTCGACACCSINENITVPDTCVKNFYAMKRMEEVGGOA	60					
1	APPLICDSRYLEERYLLLEAKAEENITTCGACACCSINENITVPDTCVKNFYAMKRMEEVGGOA	60					

Query Match	100.0%	Score 846;	DB 23;	Length 165;
Best Local Similarity	100.0%;	Pred. No. 1,1e-86;		
Matches 165;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	1	APRRLICDSRYLEERYLLEAKKENTTTTCAAEHCINENITVPPDKVNFYMKRRENGCOA	60	

Db 1 APRRLICDSRYLERLYLLEAKEAENITTCGAHCSLNENITVPDTKNVFYAMKRMEVGQA 60
 QY 61 VEWOGIALISEAVLRGQALLVNSQWPPELQJLHVDKAVGSLSTTLRALGAQKEAIS 120
 Db 61 VEWOGIALISEAVLRGQALLVNSQWPPELQJLHVDKAVGSLSTTLRALGAQKEAIS 120
 QY 121 PPDAASAPLRTITADTFRKLFRYNSNLRGKILKLYTGECACRIGD 165
 Db 121 PPDAASAPLRTITADTFRKLFRYNSNLRGKILKLYTGECACRIGD 165

RESULT 8
 AAM53061
 ID AAM53061 standard; protein; 165 AA.
 AC AAM53061;
 XX 25-MAR-2002 (first entry)
 XX Human erythropoietin (hepo), 165 residue form.
 DE Human erythropoietin (hepo), 165 residue form.
 XX Human erythropoietin; EPO; hepo; haemostatic; red blood cell;
 KW blood disorder; anaemia; chronic renal failure; CRF; AIDS;
 KM acquired immunodeficiency syndrome; cancer chemotherapy; haemostatic;
 KM anti-HIV, antianaemic.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Disulfide-bond 7..161
 FT Modified-site 24
 FT Disulfide-bond /note= "N-glycosylated"
 FT Modified-site 29..33
 FT Disulfide-bond 38
 FT Modified-site /note= "N-glycosylated"
 FT Modified-site 83
 FT Modified-site /note= "N-glycosylated"
 FT Modified-site 126
 FT Modified-site /note= "O-glycosylated"
 XX WO200187329-A1.
 XX 22-NOV-2001.
 PD 08-MAY-2001; 2001WO-EP05187.
 PF 15-MAY-2000; 2000EP-0110355.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA Papadimitriou A;
 PI WPI; 2002-082943/11.
 DR Composition useful in the treatment of e.g. AIDS comprises an
 PT erythropoietin protein, and a multiple charged inorganic anion in a
 PT buffer -
 XX Claim 28; Fig 1; 64pp; English.
 XX The invention relates to liquid pharmaceutical compositions comprising
 CC an erythropoietin (Epo) protein, a multiple negatively charged inorganic
 CC anion in a buffer which maintains the pH of the solution from 5.5-7.0,
 CC and optionally at least one excipient. The erythropoietin used in the
 CC composition is preferably human (AAM53061 or AAM53062) a human
 CC erythropoietin variant containing additional glycosylation sites
 CC (AAM53064-AAM53107), or an erythropoietin with the C-terminal addition
 CC of a C-terminal fragment of human chorionic gonadotropin (AAM53063).
 CC Erythropoietin is a glycoprotein essential for the formation of red blood
 CC cells and is therefore useful in the treatment of blood disorders
 CC characterised by low or defective red blood cell production. The
 CC compositions of the invention can be used in the treatment and prevention
 CC of anaemia in chronic renal failure patients (CRF), AIDS (acquired

CC immunodeficiency syndrome), and/or for the treatment of cancer patients
 CC undergoing chemotherapy. Unlike prior art erythropoietin compositions,
 CC the compositions of the invention do not contain human serum albumin
 CC (thereby avoiding the possibility of viral infections and allergic
 CC reactions associated with this component), are liquid rather than
 CC lyophilisates (and therefore do not need to be reconstituted before
 CC administration), and are stable at elevated temperatures such as 25
 CC degrees Celsius and even 40 degrees Celsius, and therefore can be stored
 CC without refrigeration for prolonged periods without degradation and loss
 CC of activity. The present sequence represents the 165 residue form
 CC of human erythropoietin which is specifically claimed for use in a
 CC composition of the invention.
 XX
 SQ Sequence 165 AA:
 Query Match 100.0%; Score 846; DB 23; Length 165;
 Best Local Similarity 100.0%; Pred. No. 1,1e-86;
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 APRRLICDSRYLERLYLLEAKEAENITTCGAHCSLNENITVPDTKNVFYAMKRMEVGQA 60
 Db 1 APRRLICDSRYLERLYLLEAKEAENITTCGAHCSLNENITVPDTKNVFYAMKRMEVGQA 60
 QY 61 VEWOGIALISEAVLRGQALLVNSQWPPELQJLHVDKAVGSLSTTLRALGAQKEAIS 120
 Db 61 VEWOGIALISEAVLRGQALLVNSQWPPELQJLHVDKAVGSLSTTLRALGAQKEAIS 120
 QY 121 PPDAASAPLRTITADTFRKLFRYNSNLRGKILKLYTGECACRIGD 165
 Db 121 PPDAASAPLRTITADTFRKLFRYNSNLRGKILKLYTGECACRIGD 165

RESULT 9
 AAP70398
 ID AAP70398 standard; protein; 166 AA.
 AC AAP70398;
 XX 19-FEB-1991 (first entry)
 DT Sequence of human erythropoietin (EPO).
 XX
 DE Mega-karyocyte-platelet growth factor; hormone;
 KW mega-karyocyte colony stimulating factor; therapy;
 KW small acetyl cholinesterase positive cell;
 KW erythrocyte growth effect.
 XX Homo sapiens.
 OS
 XX JP62149624-A.
 PN 03-JUL-1987.
 PD 15-AUG-1986; 86JP-0191542.
 PF 13-SEP-1985; 85JP-0203049.
 PR 15-AUG-1986; 86JP-0191542.
 XX (KAWA/) KAWAKITA M.
 PA WPI; 1987-224837/32.
 DR Megakaryocyte-platelet growth factor - contains as active
 PT component human erythropoietin and is used to treat diseases
 PT caused by decrease in platelets
 PS Disclosure; Page 181; 8pp; Japanese.
 XX All of the Cys residues in the SQ are labelled "SH". Megakaryocyte-
 CC platelet growth factor contains human EPO as an active principle.
 CC Human EPO has a megakaryocyte colony-stimulating activity and
 CC increases the ratio of small acetyl cholinesterase positive cell
 CC (SACHe+) which is immature megakaryocyte. Human EPO effects

CC megakaryocyte-platelet system other than an erythrocyte growth
 CC effect. Megakaryocyte-platelet growth is usable as a remedy for
 CC diseases caused by a platelet decrease.
 XX

SO Sequence 166 AA;

Query Match 100.0%; Score 846; DB 8; Length 166;
 Best Local Similarity 100.0%; Pred. No. 1.1e-86;
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRRLICDSRYLERYLLEAKAEANITTCGAHCGLNENITVPDTKYVFYAKRMREYGOQA 60
 |||
 DB 1 APRRLICDSRYLERYLLEAKAEANITTCGAHCGLNENITVPDTKYVFYAKRMREYGOQA 60
 |||
 QY 61 VEWOGIALISEAVLRGQALLVNSQPWEPIQLHVDKAVSGLSRLTLLRALGAQKEAIS 120
 |||
 DB 61 VEWOGIALISEAVLRGQALLVNSQPWEPIQLHVDKAVSGLSRLTLLRALGAQKEAIS 120
 |||
 QY 121 PPDAASAAPLRTTTADTFPRKLFRRVYSNPLRGKILKLYTGEACRTGD 165
 |||
 DB 121 PPDAASAAPLRTTTADTFPRKLFRRVYSNPLRGKILKLYTGEACRTGD 165
 |||

RESULT 10

AAR3593
 ID AAR3593 standard; Protein; 166 AA.

AC AAR3593;

DT 20-OCT-1992 (first entry)

DE Recombinant hematopoietic molecule portion 2.

KW Erythropoietin; EPO; erythrocytes; IL-3; haematopoiesis.

OS Homo sapiens.

PN WO9206116-A.

PD 16-APR-1992.

PF 26-SEP-1991; 91WO-US07053.

PR 28-SEP-1990; 90US-0589958.

PA (ORTH) ORTHO PHARM CORP.

PI Rosen JT;

DR WPI: 1992-150819/18.

PT Recombinant haematopoietic molecules useful in treating
 anaemia(s) - comprise IL-3 or GM-CSF, EPO, G-CSF, IL-5 or M-CSF
 and has early and later myeloid differentiation activity

PS Disclosure; Page 32; 82pp; English.

CC This protein sequence given comprises the entire amino acid sequence
 CC of human erythropoietin (EPO). EPO leads to the maturation of
 CC erythrocytes and is therefore designated as a late myeloid
 CC differentiation factor (MDF). Within the scope of the invention
 CC hybrid molecules were produced which contain at least a portion of an
 CC early MDF and at least a portion of a late MDF covalently linked. The
 CC EPO sequence given is effective within the scope of the invention in
 CC full or in a truncated version. Amino acids 7-161 act as a
 CC late MDF when recombined with an early MDF eg. IL-3.
 CC These compounds can be used to promote haematopoiesis in a patient.
 CC The bonding of the early and late factors allows a very high conc. of
 CC late MDF at the surface of a cell which the early MDF is bound. It
 CC also allows the early MDA to act more specifically to stimulate only
 CC the desired lineage, thus reducing undesirable effects. These
 CC compounds are useful for treating anaemias of various origins eg. renal
 CC failure and AIDS. It is easier to produce and administer one

CC recombinant molecule rather than two separate molecules.

SO Sequence 166 AA;

Query Match 100.0%; Score 846; DB 13; Length 166;
 Best Local Similarity 100.0%; Pred. No. 1.1e-86;
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRRLICDSRYLERYLLEAKAEANITTCGAHCGLNENITVPDTKYVFYAKRMREYGOQA 60
 |||
 DB 1 APRRLICDSRYLERYLLEAKAEANITTCGAHCGLNENITVPDTKYVFYAKRMREYGOQA 60
 |||
 QY 61 VEWOGIALISEAVLRGQALLVNSQPWEPIQLHVDKAVSGLSRLTLLRALGAQKEAIS 120
 |||
 DB 61 VEWOGIALISEAVLRGQALLVNSQPWEPIQLHVDKAVSGLSRLTLLRALGAQKEAIS 120
 |||
 QY 121 PPDAASAAPLRTTTADTFPRKLFRRVYSNPLRGKILKLYTGEACRTGD 165
 |||
 DB 121 PPDAASAAPLRTTTADTFPRKLFRRVYSNPLRGKILKLYTGEACRTGD 165
 |||

RESULT 11

AAW7780
 ID AAW7780 standard; Protein; 166 AA.

AC AAW7780;

DT 24-NOV-1998 (first entry)

DE Human EPO receptor agonist polypeptide.

KW Haematopoietic receptor agonist; erythropoietin receptor agonist;

KW EPO; human; chimeric protein; stem cell expansion; tumour;

KW infection; autoimmune disease; haematopoietic disorder; therapy;

KW dendritic cell.

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

FT	Misc-difference	37..38	/note="possible positions of new C- and N-termini"
FT	Misc-difference	38..39	/note="possible positions of new C- and N-termini"
FT	Misc-difference	39..40	/note="possible positions of new C- and N-termini"
FT	Misc-difference	40..41	/note="possible positions of new C- and N-termini"
FT	Misc-difference	41..42	/note="possible positions of new C- and N-termini"
FT	Misc-difference	42..43	/note="possible positions of new C- and N-termini"
FT	Misc-difference	43..44	/note="possible positions of new C- and N-termini"
FT	Misc-difference	44..45	/note="possible positions of new C- and N-termini"
FT	Misc-difference	45..46	/note="possible positions of new C- and N-termini"
FT	Misc-difference	46..47	/note="possible positions of new C- and N-termini"
FT	Misc-difference	47..48	/note="possible positions of new C- and N-termini"
FT	Misc-difference	48..49	/note="possible positions of new C- and N-termini"
FT	Misc-difference	49..50	/note="possible positions of new C- and N-termini"
FT	Misc-difference	50..51	/note="possible positions of new C- and N-termini"
FT	Misc-difference	51..52	/note="possible positions of new C- and N-termini"
FT	Misc-difference	52..53	/note="possible positions of new C- and N-termini"
FT	Misc-difference	53..54	/note="possible positions of new C- and N-termini"
FT	Misc-difference	54..55	/note="possible positions of new C- and N-termini"
FT	Misc-difference	55..56	/note="possible positions of new C- and N-termini"
FT	Misc-difference	56..57	/note="possible positions of new C- and N-termini"
FT	Misc-difference	57..58	/note="possible positions of new C- and N-termini"
FT	Misc-difference	77..78	/note="possible positions of new C- and N-termini"
FT	Misc-difference	78..79	/note="possible positions of new C- and N-termini"
FT	Misc-difference	79..80	/note="possible positions of new C- and N-termini"
FT	Misc-difference	81..82	/note="possible positions of new C- and N-termini"
FT	Misc-difference	82..83	/note="possible positions of new C- and N-termini"
FT	Misc-difference	84..85	/note="possible positions of new C- and N-termini"
FT	Misc-difference	85..86	/note="possible positions of new C- and N-termini"
FT	Misc-difference	86..87	/note="possible positions of new C- and N-termini"
FT	Misc-difference	87..88	/note="possible positions of new C- and N-termini"
FT	Misc-difference	88..89	/note="possible positions of new C- and N-termini"
FT	Misc-difference	108..109	/note="possible positions of new C- and N-termini"
FT	Misc-difference	109..110	/note="possible positions of new C- and N-termini"
FT	Misc-difference	110..111	/note="possible positions of new C- and N-termini"
FT	Misc-difference	111..112	/note="possible positions of new C- and N-termini"
FT	Misc-difference	112..113	/note="possible positions of new C- and N-termini"
FT	Misc-difference	113..114	/note="possible positions of new C- and N-termini"

FT	Misc-difference	114..115	/note="possible positions of new C- and N-termini"
FT	Misc-difference	115..116	/note="possible positions of new C- and N-termini"
FT	Misc-difference	116..117	/note="possible positions of new C- and N-termini"
FT	Misc-difference	117..118	/note="possible positions of new C- and N-termini"
FT	Misc-difference	118..119	/note="possible positions of new C- and N-termini"
FT	Misc-difference	119..120	/note="possible positions of new C- and N-termini"
FT	Misc-difference	120..121	/note="possible positions of new C- and N-termini"
FT	Misc-difference	121..122	/note="possible positions of new C- and N-termini"
FT	Misc-difference	122..123	/note="possible positions of new C- and N-termini"
FT	Misc-difference	123..124	/note="possible positions of new C- and N-termini"
FT	Misc-difference	124..125	/note="possible positions of new C- and N-termini"
FT	Misc-difference	125..126	/note="possible positions of new C- and N-termini"
FT	Misc-difference	126..127	/note="possible positions of new C- and N-termini"
FT	Misc-difference	127..128	/note="possible positions of new C- and N-termini"
FT	Misc-difference	128..129	/note="possible positions of new C- and N-termini"
FT	Misc-difference	129..130	/note="possible positions of new C- and N-termini"
FT	Misc-difference	130..131	/note="possible positions of new C- and N-termini"
FT	Misc-difference	131..132	/note="possible positions of new C- and N-termini"
PN	WO9817810-A2.		
PD	30-APR-1998.		
PP	23-OCT-1997; 97WO-US20037.		
PR	25-OCT-1996; 96GS-0029629.		
PA	(SEAR) SEARLE & CO G D.		
PI	Feng Y, McKearn JP, McWhetter CA, Minnerly JC, Minster NT;		
PI	Staten NR, Streeter PR, Summers NL, Woulfe SL;		
DR	WPI; 1998-261504/23.		
XX			
PT	Multi-functional chimeric haematopoietic receptor agonist - useful		
PT	to treat haematopoietic disorders, tumours, infections or autoimmune		
XX	diseases		
XX			
PS	Claim 1, Page 762; 841pp; English.		
CC	A human erythropoietin (EPO) receptor agonist polypeptide comprises		
CC	a modified EPO amino acid sequence of the formula provided in		
CC	AAW77780, in which the N-terminus is joined to the C-terminus directly		
CC	or via a linker, the polypeptide having new C- and N-termini at one		
CC	of the positions indicated. Novel claimed multi-functional chimeric		
CC	haematopoietic receptor agonists (see AAW77812-22) have the formula		
CC	R1-R1-R2, R2-R1-R1, R1-R2 or R2-R1, where L is a linker and R1 and		
CC	R2 are independently selected from: (a) the human EPO receptor		
CC	agonist; (b) a human stem cell factor receptor agonist polypeptide		
CC	(see AAW77781); (c) a human flt-3 receptor agonist polypeptide (see		
CC	AAW77782); (d) a modified human granulocyte colony stimulating factor		
CC	(G-CSF) polypeptide (see AAW77783); (e) modified human interleukin-3		
CC	polypeptide (see AAW77784); (f) modified human c-mpl ligand polypeptide		
CC	(see AAW77785); and (g) a factor selected from the group consisting of		

CC a CSF, a cytokine, a lymphokine, an interleukin and a haematopoietic
CC growth factor provided that at least R1 or R2 is selected from (a),
CC (b) or (c) as above. The multi functional chimERIC haematopoietic
CC receptor agonist can be used to stimulate the production of

Query Match
Best Local Similarity 100.0%; Score 846; DB 19; Length 166;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEKRLLEKAEKENTTGCAGHCSLNENITVPDTKYNFYAMKREYVGOQA 60
Db 1 APPRLICDSRVLEKRLLEKAEKENTTGCAGHCSLNENITVPDTKYNFYAMKREYVGOQA 60

OY 61 VEWOGIALLSFAVLRGQALLVNSQPWEPLOLHVDAVSGRLSTTLIRALGAOKEAIS 120
Db 61 VEWOGIALLSFAVLRGQALLVNSQPWEPLOLHVDAVSGRLSTTLIRALGAOKEAIS 120

OY 121 PPDASAAPLRTITADTFRKLFRVYSNPLRGKRLKLTGECACRTGD 165
Db 121 PPDASAAPLRTITADTFRKLFRVYSNPLRGKRLKLTGECACRTGD 165

RESULT 12
AAM58404
ID AAM58404 standard; Protein; 166 AA.
XX AAM58404;
AC
XX
DT 12-OCT-1998 (first entry)
XX
DE Human erythropoietin.
XX
KW Erythropoietin receptor agonist; EPO; human; anaemia;
KW haematopoietic deficiency; red blood cell; erythroid progenitor;
KW bone marrow suppression.
XX
OS Homo sapiens.
XX
PN WO9818926-A1.
XX
PD 07-MAY-1998.
XX
PF 23-OCT-1997; 97WO-US18703.
XX
PR 25-OCT-1996; 96US-0034044.
XX
PA (SEAR) SEARLE & CO G D.
XX
FI Feng Y, McWhirter CA, Summers N;
XX WPI; 1998-272221/24.
XX N-PSDB; AAV31031.
XX
PR Human erythropoietin receptor agonist polypeptide - used to
XX stimulate the production of red blood cells in a patient
XX
XX Claim 1; Page 93; 112pp; English.

A claimed human erythropoietin (EPO) receptor agonist polypeptide
CC comprises a modified EPO amino acid sequence given in AAM58404,
CC where (a) optionally 1-6 amino acids from the N-terminus and 1-5
CC of the C-terminus can be deleted, (b) the N-terminus is joined to
CC the C-terminus directly or through a linker (see AAM58405-12) capable
CC of joining the N-terminus to the C-terminus, (c) there are new C-
CC 23 and N-termini at any two consecutive amino acids from amino acids
CC 77-78 to 82-83, 84-85 to 86-85, and 108-109 to 131-132, and (d)
CC optionally the agonist polypeptide is preceded by Met, Ala, or
CC MetCatalan13072, these circularly permuted EPO receptor agonists
CC (see AAM58413-72) are claimed. Also claimed are: nucleic acid
CC methods of producing an EPO receptor agonist using transformed or
CC transfected host cells; and methods for stimulating the production

CC of haematopoietic cells, for selective ex vivo expansion of
CC erythroid progenitors, and treating patients having a haematopoietic
CC disorder using the EPO receptor agonists. The EPO receptor
CC agonists retain one or more activities of native EPO and may also
CC show improved haematopoietic cell-stimulating activity and/or an
CC improved activity profile which may include reduction of undesirable
CC biological activities associated with native EPO and/or have
CC improved physical properties such as increased solubility,
CC stability and refold efficiency.

Sequence 166 AA;
SQ

Query Match
Best Local Similarity 100.0%; Score 846; DB 19; Length 166;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEKRLLEKAEKENTTGCAGHCSLNENITVPDTKYNFYAMKREYVGOQA 60
Db 1 APPRLICDSRVLEKRLLEKAEKENTTGCAGHCSLNENITVPDTKYNFYAMKREYVGOQA 60

OY 61 VEWOGIALLSFAVLRGQALLVNSQPWEPLOLHVDAVSGRLSTTLIRALGAOKEAIS 120
Db 61 VEWOGIALLSFAVLRGQALLVNSQPWEPLOLHVDAVSGRLSTTLIRALGAOKEAIS 120

OY 121 PPDASAAPLRTITADTFRKLFRVYSNPLRGKRLKLTGECACRTGD 165
Db 121 PPDASAAPLRTITADTFRKLFRVYSNPLRGKRLKLTGECACRTGD 165

RESULT 13
ABB07030
ID ABB07030 standard; Protein; 166 AA.
XX ABB07030;
AC
XX
DT 21-JUN-2002 (first entry)
XX
DE Modified erythropoietin related gene protein sequence.
XX
KW Modified erythropoietin; EPO.
XX
OS unidentified.
XX
PN KR145802-B1.
XX
PD 01-AUG-1998.
XX
PF 31-MAY-1994; 94KR-0012082.
XX
PR 31-MAY-1994; 94KR-0012082.
XX
PA (GLDS) LG CHEM CO LTD.
XX
FI Kim C, Song Y, Lee T;
XX
DR WPI; 2000-234250/20.
XX N-PSDB; ABL50878.
XX
PR MODIFIED ERYTHROPOIETIN GENE AND EXPRESSION VECTORS THEREOF -
XX
XX Disclosure; Page 14; 15pp; Korean.

The present invention describes modified erythropoietin (EPO) genes
CC and expression vectors comprising the genes. The present sequence
CC represents a protein sequence from the present invention.
XX
XX Sequence 166 AA;
SQ

Query Match
Best Local Similarity 100.0%; Score 846; DB 21; Length 166;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEKRLLEKAEKENTTGCAGHCSLNENITVPDTKYNFYAMKREYVGOQA 60

```
Db 1 APRRLICDSRVLEKRLLEKAEENITTCGAHCSLNENITVPDTKNVFAKRMVEYGOQA 60
|||||
QY 61 VEWOGIALLSEAVLNGOALLVNSSQPWEPLQHLVDKAVSGLSLTLTLRALGAQKEAIS 120
|||||
Db 61 VEWOGIALLSEAVLNGOALLVNSSQPWEPLQHLVDKAVSGLSLTLTLRALGAQKEAIS 120
|||||
QY 121 PPDAAASAPLRTTTADTFKRLFRVYSNFLRGKILKLTGTGACRGTGD 165
|||||
Db 121 PPDAAASAPLRTTTADTFKRLFRVYSNFLRGKILKLTGTGACRGTGD 165
|||||

RESULT 14
AAE02641
ID AAE02641 standard; Protein: 166 AA.
XX
XX AAE02641;
XX
XX
XX 06-AUG-2001 (first entry)
XX
XX Human erythropoietin (EPO) mature protein.
DE
XX Human erythropoietin; EPO; antianaemic; nephroretrophic; anti-HIV;
KW vaccine; haemostatic; immunoglobulin; Ig; EPO deficient disease;
KW anaemia; renal failure; Human Immunodeficiency Virus; HIV;
KW haematopoietic growth factor.
XX
XX Homo sapiens.
OS
XX WO200136489-A2.
PN
XX 25-MAY-2001.
PD
XX 03-NOV-2000; 2000WO-EP10843.
PF
XX
XX 12-NOV-1999; 99US-0164855.
PR
XX (MERE ) MERCK PATENT GMBH.
PA
XX Hartmann A, Brandt S, Rieke E, Sobel C, Lo K, Way JC, Gillies S;
PI
XX WPI; 2001-367563/38.
DR
XX N-PSDB; AAD06893.
DR
XX
XX Novel modified erythropoietin forms such as fusion proteins, comprising
PT
XX the biological activity of erythropoietin forms -
PT
XX
XX Example 1; Page 22; 58pp; English.
PS
XX
XX The present sequence is human erythropoietin (EPO) mature protein. EPO
CC
XX has improved biological activity and an extended serum half life greater
CC
XX than 20 hours. The present invention relates to modified EPO forms such
CC
XX as fusion proteins comprising a FC portion of an immunoglobulin (Ig)
CC
XX molecule and an EPO molecule (Fc-EPO). The FC portion is fused covalently
CC
XX through its C-terminus directly or indirectly to the EPO molecule, and
CC
XX where the FC portion as well as EPO portion may be modified or mutated.
CC
XX The invention also relates to non-fused EPO molecules which have a
CC
XX pattern of cysteines or disulphide bonding which is distinct from human
CC
XX or animal EPO. Pharmaceutical compositions containing EPO are useful in
CC
XX the treatment of EPO deficient diseases such as anaemia, renal failure,
CC
XX HIV infection, blood loss and chronic disease that can be treated with
CC
XX haematopoietic growth factor.
XX
XX
XX Sequence 166 AA;
SQ
Query Match 100.0%; Score 846; DB 22; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 APRRLICDSRVLEKRLLEKAEENITTCGAHCSLNENITVPDTKNVFAKRMVEYGOQA 60
|||||
Db 1 APRRLICDSRVLEKRLLEKAEENITTCGAHCSLNENITVPDTKNVFAKRMVEYGOQA 60
|||||
```

```
QY 61 VEWOGIALLSEAVLNGOALLVNSSQPWEPLQHLVDKAVSGLSLTLTLRALGAQKEAIS 120
|||||
Db 61 VEWOGIALLSEAVLNGOALLVNSSQPWEPLQHLVDKAVSGLSLTLTLRALGAQKEAIS 120
|||||
QY 121 PPDAAASAPLRTTTADTFKRLFRVYSNFLRGKILKLTGTGACRGTGD 165
|||||
Db 121 PPDAAASAPLRTTTADTFKRLFRVYSNFLRGKILKLTGTGACRGTGD 165
|||||

RESULT 15
AAB66698
ID AAB66698 standard; protein: 166 AA.
XX
XX AAB66698;
XX
XX 06-APR-2001 (first entry)
XX
XX Human erythropoietin protein #2.
DE
XX
XX Erythropoietin; EPO; reticulocytes; red blood cell;
KW ethylene glycol; chronic renal failure; AIDS; cancer.
XX
XX Homo sapiens.
OS
XX WO200102017-A2.
PN
XX 11-JAN-2001.
PD
XX 28-JUN-2000; 2000WO-EP06009.
PF
XX 02-JUL-1999; 99US-0142243.
PR
XX 05-AUG-1999; 99US-0147452.
PR
XX 30-AUG-1999; 99US-0151454.
XX
XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA
XX Burg J, Hilger B, Josel H;
PI
XX WPI; 2001-147051/15.
DR
XX
XX Novel erythropoietin-glycoprotein conjugate useful for treating
PT
XX diseases correlated with anemia in chronic renal failure patients, AIDS
PT
XX and for treating cancer, is linked to polyethylene glycol through
PT
XX linker -
PT
XX
XX Claim 19; Fig 2; 40pp; English.
PS
XX
XX The present invention relates to a conjugate comprising, human
CC
XX erythropoietin glycoprotein (EPO) having at least one free
CC
XX amino group and having in vivo biological activity of causing
CC
XX an increase the production of reticulocytes and red blood cells,
CC
XX covalently linked to 1-3 lower-alkoxy poly(ethylene glycol) groups
CC
XX through a linker. The invention is useful for preparation of
CC
XX medicaments for the treatment of prophylaxis of disease correlated
CC
XX with anemia in chronic renal failure patients (CRF), AIDS and for
CC
XX the treatment of cancer patients undergoing chemotherapy.
XX
XX
XX Sequence 166 AA;
SQ
Query Match 100.0%; Score 846; DB 22; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 APRRLICDSRVLEKRLLEKAEENITTCGAHCSLNENITVPDTKNVFAKRMVEYGOQA 60
|||||
Db 1 APRRLICDSRVLEKRLLEKAEENITTCGAHCSLNENITVPDTKNVFAKRMVEYGOQA 60
|||||
QY 61 VEWOGIALLSEAVLNGOALLVNSSQPWEPLQHLVDKAVSGLSLTLTLRALGAQKEAIS 120
|||||
Db 61 VEWOGIALLSEAVLNGOALLVNSSQPWEPLQHLVDKAVSGLSLTLTLRALGAQKEAIS 120
|||||
QY 121 PPDAAASAPLRTTTADTFKRLFRVYSNFLRGKILKLTGTGACRGTGD 165
|||||
Db 121 PPDAAASAPLRTTTADTFKRLFRVYSNFLRGKILKLTGTGACRGTGD 165
|||||
```

Fri Jan 3 10:09:17 2003

us-09-853-731-1.rag

Page 10

Db 121 PDDASAPLRTTTADPFKKLFRYNSNLRGKLYTGACRTGD 165

Search completed: January 2, 2003, 15:11:27
Job time : 32.9033 secs


```

; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 165
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-014-363-1

Query Match
Best Local Similarity 100.0%; Score 846; DB 12; Length 165;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 165
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-014-363-1

Query Match
Best Local Similarity 100.0%; Score 846; DB 12; Length 165;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 165
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-014-363-1

Query Match
Best Local Similarity 100.0%; Score 846; DB 12; Length 165;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 165
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-014-363-1

Query Match
Best Local Similarity 100.0%; Score 846; DB 12; Length 165;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 166
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-014-363-2

Query Match
Best Local Similarity 100.0%; Score 846; DB 12; Length 166;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 169
; TYPE: PRF
; ORGANISM: CHO/dhfr-
US-10-014-363-4

Query Match
Best Local Similarity 100.0%; Score 846; DB 12; Length 169;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 169
; TYPE: PRF
; ORGANISM: CHO/dhfr-
US-10-014-363-4

Query Match
Best Local Similarity 100.0%; Score 846; DB 12; Length 169;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
Qy 61 VEVWOGIALLSEAVLRGQALLVNSSQPWEPIQLHVDKAVSGLSLTTLLRALGAQKEAIS 120
|
Db 64 VEVWOGIALLSEAVLRGQALLVNSSQPWEPIQLHVDKAVSGLSLTTLLRALGAQKEAIS 123
|
Qy 121 PPDAASAPLRTITADTFRKLFRVYSNPLRGKLLKLTGECRTGD 165
|
Db 124 PPDAASAPLRTITADTFRKLFRVYSNPLRGKLLKLTGECRTGD 168

RESULT 6
US-10-014-363-3
; Sequence 3, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 174
; TYPE: PRT
; ORGANISM: CHO/dhfr-
US-10-014-363-3

Query Match 100.0%; Score 846; DB 12; Length 174;
Best Local Similarity 100.0%; Pred. No. 3e-84;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APRRICSRVLERLLEAKAEENITTCACHECSINENITVPDTRKVNRYAMKRMVGOOA 60
|
Db 9 APRRICSRVLERLLEAKAEENITTCACHECSINENITVPDTRKVNRYAMKRMVGOOA 68
|
Qy 61 VEVWOGIALLSEAVLRGQALLVNSSQPWEPIQLHVDKAVSGLSLTTLLRALGAQKEAIS 120
|
Db 69 VEVWOGIALLSEAVLRGQALLVNSSQPWEPIQLHVDKAVSGLSLTTLLRALGAQKEAIS 128
|
Qy 121 PPDAASAPLRTITADTFRKLFRVYSNPLRGKLLKLTGECRTGD 165
|
Db 129 PPDAASAPLRTITADTFRKLFRVYSNPLRGKLLKLTGECRTGD 173

RESULT 7
US-10-014-363-5
; Sequence 5, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 174
; TYPE: PRT
; ORGANISM: CHO/dhfr-
US-10-014-363-5

Query Match 100.0%; Score 846; DB 12; Length 174;
Best Local Similarity 100.0%; Pred. No. 3e-84;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APRRICSRVLERLLEAKAEENITTCACHECSINENITVPDTRKVNRYAMKRMVGOOA 60
|
Db 9 APRRICSRVLERLLEAKAEENITTCACHECSINENITVPDTRKVNRYAMKRMVGOOA 68
|
Qy 61 VEVWOGIALLSEAVLRGQALLVNSSQPWEPIQLHVDKAVSGLSLTTLLRALGAQKEAIS 120
|
Db 69 VEVWOGIALLSEAVLRGQALLVNSSQPWEPIQLHVDKAVSGLSLTTLLRALGAQKEAIS 128
|
Qy 121 PPDAASAPLRTITADTFRKLFRVYSNPLRGKLLKLTGECRTGD 165
|
Db 129 PPDAASAPLRTITADTFRKLFRVYSNPLRGKLLKLTGECRTGD 173

RESULT 8
US-09-864-761-48502
; Sequence 48502, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Rank, David R.
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecm1ca-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48502
; LENGTH: 60
; TYPE: PRT
```

ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AF053356.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.93
OTHER INFORMATION: SWISSPROT HIT: P01588, EVALUO 9.00e-28
OTHER INFORMATION: EST_HUMAN HIT: AA662379.1, EVALUO 3.00e-10
US-09-864-761-48502

Query Match 34.8%; Score 294; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.4e-25;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 56 VGGQAVVWQGLALSEAVLRQALVNSQPEPLQHVRAVGLSLTTLRALGQ 115
DB 1 VGGQAVVWQGLALSEAVLRQALVNSQPEPLQHVRAVGLSLTTLRALGQ 60

RESULT 9
US-09-975-063-2
Sequence 2, Application US/09975063
Patent No. US20020045255A1
GENERAL INFORMATION:

APPLICANT: POWELL, Jerry S.
TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE, HIGH LEVEL
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: Dorsey & Whitney, LLP
STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.7/8
CURRENT APPLICATION DATA: US/09/975.063
FILING DATE: 10-03-2001
ATTORNEY/AGENT INFORMATION:
NAME: Robert's, Mark W.
REGISTRATION NUMBER: 46,160
REFERENCE/DOCKET NUMBER: 500582.03 (112893.109)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-903-8728
TELEFAX: 206-903-8820

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2

US-09-975-063-2

Query Match 13.1%; Score 111; DB 10; Length 30;
Best Local Similarity 83.3%; Pred. No. 8e-06;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 APPRLIDSRVLEKYLEAKAEENITGCA 30
DB 1 APXRLIDSRVLEKYLEAKAEEXITDGA 30

RESULT 10
US-10-011-858-2
Sequence 2, Application US/10011858
Patent No. US20020137145A1
GENERAL INFORMATION:
APPLICANT: POWELL, Jerry S.

TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE, HIGH LEVEL
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: Dorsey & Whitney, LLP
STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-WORD
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/011,858
FILING DATE: 05-NO. US20020137145A1-2001
ATTORNEY/AGENT INFORMATION:
NAME: Robert's, Mark W.
REGISTRATION NUMBER: 46,160
REFERENCE/DOCKET NUMBER: 500582.13
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-903-8728
TELEFAX: 206-903-8820

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2

US-10-011-858-2

Query Match 13.1%; Score 111; DB 12; Length 30;
Best Local Similarity 83.3%; Pred. No. 8e-06;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 APPRLIDSRVLEKYLEAKAEENITGCA 30
DB 1 APXRLIDSRVLEKYLEAKAEEXITDGA 30

RESULT 11
US-09-919-703-9
Sequence 9, Application US/09919703
Patent No. US20020165129A1
GENERAL INFORMATION:

APPLICANT: Krystal, Gerald
TITLE OF INVENTION: Peptides and Their Use to Ameliorate
Cell Death
FILE REFERENCE: 50216/003004
CURRENT APPLICATION NUMBER: US/09/919,703
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 09/294,457
PRIOR FILING DATE: 1999-04-19
PRIOR APPLICATION NUMBER: US 08/759,599
PRIOR FILING DATE: 1996-12-05
PRIOR APPLICATION NUMBER: US 60/008,233
PRIOR FILING DATE: 1995-12-06
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9

LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic polypeptide
US-09-919-703-9

Query Match 12.2%; Score 103; DB 9; Length 20;

Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0;

QY 147 NFLRGKLYTGACRTGD 165

Db 1 NFLRGKLYTGACRTGD 19

RESULT 12

US-09-919-703-10

; Sequence 10, Application US/09919703

; Patent No. US20020165129A1

; GENERAL INFORMATION:

; APPLICANT: Krystal, Gerald

; APPLICANT: Rabbitt, Simon W.

; TITLE OF INVENTION: Peptides and Their Use to Ameliorate

; FILE REFERENCE: 50216/003004

; CURRENT APPLICATION NUMBER: US/09/919,703

; CURRENT FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: US 09/294,457

; PRIOR FILING DATE: 1999-04-19

; PRIOR APPLICATION NUMBER: US 08/759,599

; PRIOR FILING DATE: 1996-12-05

; PRIOR APPLICATION NUMBER: US 60/008,233

; PRIOR FILING DATE: 1995-12-06

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 20

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic polypeptide

US-09-919-703-10

Query Match 10.4%; Score 88; DB 9; Length 20;

Best Local Similarity 95.0%; Pred. No. 0.0014;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RLICDSRVLEARYLLEAKAE 23

Db 1 RLICDSRVLEARYLLEAKAE 20

RESULT 13

US-09-801-368-244

; Sequence 244, Application US/09801368

; Patent No. US20020128250A1

; GENERAL INFORMATION:

; APPLICANT: Busby, Robert

; APPLICANT: Cail, Brian

; APPLICANT: Hecht, Peter

; APPLICANT: Holtzman, Doug

; APPLICANT: Madden, Kevin

; APPLICANT: Maxon, Mary

; APPLICANT: Milne, Todd

; APPLICANT: No. US20020128250A1man, Thea

; APPLICANT: Royer, John

; APPLICANT: Salama, Sofie

; APPLICANT: Sherman, Amir

; APPLICANT: Silva, Jeff

; APPLICANT: Summers, Eric

; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi

; FILE REFERENCE: 109272.147

; CURRENT APPLICATION NUMBER: US/09/801,368

; CURRENT FILING DATE: 2001-03-07

; PRIOR APPLICATION NUMBER: US 09/487,558

; PRIOR FILING DATE: 2000-01-19

; PRIOR APPLICATION NUMBER: US 60/160,587

; PRIOR FILING DATE: 1999-10-20

; NUMBER OF SEQ ID NOS: 440

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 244

; LENGTH: 1564

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

US-09-801-368-244

Query Match 8.7%; Score 74; DB 10; Length 1564;
Best Local Similarity 21.7%; Pred. No. 21;

Matches 31; Conservative 15; Mismatches 51; Indels 46; Gaps 4;

QY 39 ITVPDTKYNFVAMKRMVGOQ-----AVEVWGIALLSAVALRGOA 79

Db 1057 VGVPGGVLNVEQRKRLTIVELAAKRLVFLDEPTSGLDSTAMSTCOLMKKLASRGOA 1116

QY 80 LLYNSSQWPEPIQLHDKAV-----SGLRSLTTLRALCAQKEAISPPDA 124

Db 1117 ILCTTHQPSALLMOEDRLLFTQEGQGVYFGEIGKCKTMINYFBAHGANH---CPPDA 1173

QY 125 -----ASAAPLRTITADTF 138

Db 1174 NPDEMLETIVGAAPGTHASDYF 1196

RESULT 14

US-09-764-877-1947

; Sequence 1947, Application US/09764877

; Patent No. US20020147140A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC005

; CURRENT APPLICATION NUMBER: US/09/764,877

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 4031

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1947

; LENGTH: 81

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-764-877-1947

Query Match 8.6%; Score 72.5; DB 10; Length 81;
Best Local Similarity 33.7%; Pred. No. 0.48;

Matches 28; Conservative 7; Mismatches 37; Indels 11; Gaps 4;

QY 51 WKRMVGOQAVEVWGIALLSAVALRGOALLVSSQPW---EPLQLHVDKAVSGLRSIT 106

Db 6 WDRMEAAASSSLVSGPALLCSGVSAGAS--VELSLPMFSPPRGLKTPPKVYASL----- 59

QY 107 TLLRALCAQKEAISPPDAASAP 129

Db 60 SVVRHLCNQPEAL-PRGTTSTVP 81

RESULT 15

US-09-945-182-26

; Sequence 26, Application US/09945182

; Patent No. US20020160494A1

; GENERAL INFORMATION:

; APPLICANT: Celeste, Anthony J.

; APPLICANT: Wozney, John

; APPLICANT: Rosen, Vicki A.

; APPLICANT: Wolfman, Neil

; APPLICANT: Thomsen, Gerald H.

; APPLICANT: Melton, Douglas A.

; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS

; NUMBER OF SEQUENCES: 35

; CORRESPONDENCE ADDRESS:

ADDRESS: GENETICS INSTITUTE, INC.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: Massachusetts

Fri Jan 3 10:09:17 2003

us-09-853-731-1.rapb

Page 6

COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA: Release #1.0, Version #1.25
APPLICATION NUMBER: US/09/945,182
FILING DATE: 31-Aug-2001
CLASSIFICATION: <unknown>
Prior APPLICATION DATA:
APPLICATION NUMBER: 08/808,324
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5202-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5851
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-945-182-26

Query Match 8.5%; Score 72; DB 9; Length 321;
Best Local Similarity 27.2%; Pred. No. 3.8; Mismatches 48; Indels 24; Gaps 3;
Matches 31; Conservative 11;
OY 57 GQQAWEVWQGLALSEAVLKGALLVNSSQWPEPLQIHVDKAVSGLSRLTLLRALGAQK 116
DB 77 GWEVFPVWQGL-----RHQPMKOLCLELRAAWGELDAGEAERARAGPQ 120
OY 117 EAISPPDAASAPLRTITADYFKLFPRVYSNFLRGKL-----KIYTGACRTG 164
DB 121 P--PPDLRLSLGFGRRVRPPQERALLVFTRSQRKNTLFAEMREOLGSAEAGPG 172

Search completed: January 2, 2003, 15:13:49
Job time : 9.4/432 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:10:05 ; Search time 13.9577 Seconds

(without alignments)

1136.446 Million cell updates/sec

Title: US-09-853-731-1

Sequence: 1 APRRLICDSVLEKRYLLEAK.....SNFLRGRKLYTGECARTGD 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	846	100.0	193	1 ZUHU	erythropoietin pre
2	764.5	90.4	192	1 JQ0173	erythropoietin pre
3	759.5	89.8	192	1 I84613	erythropoietin pre
4	713	84.3	188	1 I46083	erythropoietin pre
5	701	82.9	192	1 S28148	erythropoietin pre
6	685.5	81.0	194	1 I46401	erythropoietin pre
7	681	80.5	192	1 A24902	erythropoietin pre
8	680.5	80.4	195	1 UC7699	erythropoietin pre
9	678	80.1	190	2 I46578	erythropoietin - r
10	638	75.4	175	2 I46199	erythropoietin - p
11	602.7	90.6	353	2 G02729	erythropoietin - d
12	89	10.5	353	2 I80105	thrombopoietin - h
13	88	10.4	323	2 AB0323	thrombopoietin pre
14	87.5	10.3	346	2 AE0959	ribonucleoside-dip
15	86	10.2	286	2 A55530	Solute binding rec
16	83	9.8	296	2 A10443	megakaryocyte grow
17	83	9.8	339	2 A83274	probable 2-hydroxy
18	80.5	9.5	3033	1 GNMVJ8	UDP-N-acetylpyruvo
19	79.5	9.4	1829	2 T35681	genome polyprotein
20	79	9.3	480	2 S56639	probable sensory h
21	78.5	9.3	813	2 AF0526	ribosomal protein
22	78.5	9.3	897	2 A54656	ATP-dependent heli
23	78	9.2	348	2 T35450	EGF receptor subst
24	78	9.2	455	2 H97693	ABC transporter AT
25	78	9.2	455	2 AG2919	methylaniline utiliz
26	77.5	9.2	747	1 S36741	conserved hypotnet
27	77.5	9.2	242	2 AD1928	probable copper-tr
28	77	9.1	451	2 S75569	hypothetical prote
29	76.5	9.0	154	2 H82810	bacterioferritin x

30	76.5	9.0	425	2 AE3465	mandelate racemase
31	75.5	8.9	637	2 S75772	hypothetical prote
32	74.5	8.8	400	2 AB2922	conserved hypotnet
33	74.5	8.8	425	2 C97696	rits beta (AF305057
34	74.5	8.8	824	2 D64738	ATP-dependent heli
35	74	8.7	282	2 B37994	RF2 protein - salm
36	74	8.7	326	2 JC4125	thrombopoietin pre
37	74	8.7	335	2 AH3625	ribonucleoside-dip
38	74	8.7	1564	2 S55517	probable transport
39	73.5	8.7	401	2 H83911	hypothetical prote
40	73.5	8.7	476	1 S71789	GCN5 protein - hum
41	73.5	8.7	717	2 F82613	VaCB protein XFl98
42	73	8.6	263	2 B75361	WD-repeat family p
43	73	8.6	1089	2 S53978	PSF1 protein - yea
44	72.5	8.6	379	2 H69478	NADH2 dehydrogenas
45	72.5	8.6	401	2 AF3341	precorrin-6y c5,15

ALIGNMENTS

RESULT 1

ZUHU

erythropoietin precursor [validated] - human

C:Species: Homo sapiens (man)

C>Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 08-Dec-2000

C:Accession: A01855; A24744; A25384; A22210; S56178

R:Jacobs, K.; Shoemaker, C.; Ruderstorf, R.; Neill, S.D.; Kaufman, R.J.; Mufson, A.;

Nature 313, 806-810, 1985

A:Title: Isolation and characterization of genomic and cDNA clones of human erythro

A:Reference number: A01855; MUID:85137899; PMID:3838366

A:Accession: A01855

A:Molecule type: mRNA; DNA

A:Residues: 1-193 <JUC>

A:Cross-references: GB:X02157; GB:X02158

R:Lin, F.K.; Sugars, S.; Lin, C.H.; Browne, J.K.; Smalling, R.; Egrise, J.C.; Chen, K.K

Proc. Natl. Acad. Sci. U.S.A. 82, 7580-7584, 1985

A:Title: Cloning and expression of the human erythropoietin gene.

A:Reference number: A24744; MUID:86067948; PMID:3865178

A:Accession: A24744

A:Molecule type: DNA

A:Residues: 1-193 <LIN>

A:Cross-references: GB:M11319; NID:9182197; PID:AA52400.1; PID:9182198

R:Lin, P.H.; Everett, R.; Wang, F.F.; Arakawa, T.; Goldwasser, E.

J. Biol. Chem. 261, 3116-3121, 1986

A:Title: Structural characterization of human erythropoietin.

A:Reference number: A25384; MUID:86140080; PMID:3949763

A:Accession: A25384

A:Molecule type: protein

A:Residues: 28-86, 'Q', 87-193 <LA1>

A:Note: forms without the carboxyl-terminal residue and the four carboxyl-terminal re

R:Yanagawa, S.; Hirade, K.; Ohnoka, H.; Sasaki, R.; Chiba, H.; Goto, M.

J. Biol. Chem. 259, 2707-2710, 1984

A:Title: Isolation of human erythropoietin with monoclonal antibodies.

A:Reference number: A22210; MUID:84135751; PMID:6698988

A:Accession: A22210

A:Molecule type: protein

A:Residues: 28-29, 'X', 31-33, 'L', 35-50, 'X', 52-53, 'D', 55, 'G', 57 <YAN>

R:Matsumoto, S.; Ikura, K.; Ueda, M.; Sasaki, R.

Plant Mol. Biol. 27, 1163-1172, 1995

A:Title: Characterization of a human glycoprotein (erythropoietin) produced in cultur

A:Reference number: S56178; MUID:95284365; PMID:7768897

A:Accession: S56178

A:Molecule type: protein

C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live

A:Residues: 28-33, 'X', 35-37 <MTS>

C:Genetics:

A:Gene: GDB:EPO

A:Cross-references: GDB:119110; OMIM:133170

A:Map position: 7q21.3-7q22.1

A:introns: 5/1; 53/3; 82/3; 142/3

C:Function:

A:Description: the primary inducer of erythrocyte formation
C:Superfamily: erythropoietin
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-193/Product: erythropoietin #status experimental <MAT>
F:34-188,56-60/Disulfide bonds: #status experimental
F:51,65,110/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:153/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 846; DB 1; Length 193;
Best Local Similarity 100.0%; Pred. No. 1,8e-74;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLLEAKAEENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 60
DB 28 APPRLICDSRVLEERYLLLEAKAEENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 87
OY 61 VEWMOGLALSEAVLRGQALLVNSQPEPQLQHDVKAAGSLRSLTTLRLALGAOKRAIS 120
DB 88 VEWMOGLALSEAVLRGQALLVNSQPEPQLQHDVKAAGSLRSLTTLRLALGAOKRAIS 147
OY 121 PPDAASAAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECRTGD 165
DB 148 PPDAASAAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECRTGD 192

RESULT 2

erythropoietin precursor - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C:Date: 07-Sep-1990 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C:Accession: J00173
R:Lin, F.K.; Lin, C.H.; Lai, P.H.; Browne, J.K.; Egrle, J.C.; Smalling, R.; Fox, G.M.; C
Gene 44, 201-209, 1986
A:Title: Monkey erythropoietin gene: cloning, expression and comparison with the human
A:Reference number: J00173; MUID:87055236; PMID:2877922
A:Accession: J00173
A:Molecule type: mRNA
A:Residues: 1-192 <LIN>
A:Cross-references: GB:M18189; GB:M15818; GB:M15819; GB:M18188; NID:g342093; PIDN:AAA368
A:Experimental source: kidney
C:Comment: This protein is the principal hormone involved in the regulation of erythrocy
C:Function: Erythropoietin is produced by kidney or liver of adult mammals and by liver
A:Description: the primary inducer of erythrocyte formation
C:Superfamily: erythropoietin
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-193/Product: erythropoietin #status experimental <MAT>
F:34-188,56-60/Disulfide bonds: #status predicted
F:51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:153/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 90.4%; Score 764.5; DB 1; Length 192;
Best Local Similarity 91.5%; Pred. No. 1.4e-66;
Matches 151; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

OY 1 APPRLICDSRVLEERYLLLEAKAEENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 60
DB 28 APPRLICDSRVLEERYLLLEAKAEENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 87
OY 61 VEWMOGLALSEAVLRGQALLVNSQPEPQLQHDVKAAGSLRSLTTLRLALGAOKRAIS 120
DB 88 VEWMOGLALSEAVLRGQALLVNSQPEPQLQHDVKAAGSLRSLTTLRLALGAOKRAIS 146
OY 121 PPDAASAAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECRTGD 165
DB 147 PPDAASAAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECRTGD 191

RESULT 3
184613
erythropoietin precursor - rhesus macaque

C:Species: Macaca mulatta (rhesus macaque)

C:Date: 02-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999

C:Accession: I84613

R:Men, D.; Boissel, J.

Blood 82, 1507-1516, 1993

A:Title: Erythropoietin structure-function relationships: High degree of sequence hom

A:Reference number: I46083; MUID:93372347; PMID:8364201

A:Accession: I84613

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Cross-references: GB:I10609; NID:g342095; PIDN:AAA36842.1; PID:g342096

C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live

C:Function: the primary inducer of erythrocyte formation

A:Description: the primary inducer of erythrocyte formation

C:Superfamily: erythropoietin

C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver

F:1-27/Domain: signal sequence #status predicted <SIG>

F:28-192/Product: erythropoietin #status predicted <MAT>

F:34-187,56-60/Disulfide bonds: #status predicted

F:51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:153/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 89.8%; Score 759.5; DB 1; Length 192;
Best Local Similarity 90.3%; Pred. No. 4.3e-66;
Matches 149; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

OY 1 APPRLICDSRVLEERYLLLEAKAEENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 60
DB 28 APPRLICDSRVLEERYLLLEAKAEENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 87
OY 61 VEWMOGLALSEAVLRGQALLVNSQPEPQLQHDVKAAGSLRSLTTLRLALGAOKRAIS 120
DB 88 VEWMOGLALSEAVLRGQALLVNSQPEPQLQHDVKAAGSLRSLTTLRLALGAOKRAIS 146
OY 121 PPDAASAAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECRTGD 165
DB 147 PPDAASAAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECRTGD 191

RESULT 4

erythropoietin precursor - cat (fragment)
C:Species: Felis silvestris catus (domestic cat)
C:Date: 16-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C:Accession: I46083
R:Men, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A:Title: Erythropoietin structure-function relationships: High degree of sequence hom
A:Reference number: I46083; MUID:93372347; PMID:8364201
A:Accession: I46083
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-188 <MEN>
A:Cross-references: GB:I10606; NID:g163820; PIDN:AAA30807.1; PID:g163821
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C:Function: the primary inducer of erythrocyte formation
C:Superfamily: erythropoietin
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-22/Domain: signal sequence (fragment) #status predicted <SIG>
F:23-188/Product: erythropoietin #status experimental <MAT>
F:29-183,51-55/Disulfide bonds: #status predicted
F:46,60,105/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:148/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 84.3%; Score 713; DB 1; Length 188;
Best Local Similarity 84.2%; Pred. No. 1.3e-61;
Matches 139; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLLEAKAEENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 60
DB 23 APPRLICDSRVLEERYLLLEAKAEENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 82

Query Match	80.5%;	Score 681;	DB 1;	Length 192;
Best Local Similarity	79.4%;	Pred. No. 1.7e-58;		
Matches 131;	Conservative 14;	Mismatches 20;	Indels 0;	Gaps 0;

QY	1	APPRICCSRVLEERLLEAKEAENITTTGCAEHCSCSINENTVDTVDVNVNPAKRMVEGOOA	60
Db	27	APPRICCSRVLEERLLEAKEAENVTMGCAESPRISENITVADTVNVNPAKRMVEEDBA	86
QY	61	VEWMOGLALSEAVYRGQAALLVNSOPPEPLOTIHYDKVASCGISRTTTLTRLAGAOKEAIS	120
Db	87	IENVWGSLSEFAIIQAQALANSSQPETLOTHIDKALISGRISSTLSRLVYGAOKELMS	146
QY	121	PPDAASAPLRKIITADTFRKLFRVYSNFLRGRKLKIYTGBACRTGD	165
Db	147	PPDTPPAPLRKIIVDTDFCKLFERYVANFLRGRKLKIYTGVEVCRRGD	191

RESULT 8
JC7699
erythropoietin - rabbit
C:Species: *Oryctolagus cuniculus* (domestic rabbit)
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
C:Accession: JC7699
R:Vitalta, A.; Wu, D.; Margalith, M.; Hobart, P.
Biochem. Biophys. Res. Commun. 284, 823-827, 2001
A:Title: Rabbit Epo gene and cDNA: Expression of rabbit Epo after intramuscular injection
A:Reference number: JC7699; MUID:21290682; PMID:11396976
A:Contents: Kidney
A:Accession: JC7699
A:Model: 3D
A:Residues: 1-195
A:Cross-references: GB:AF290943
C:Comment: This protein, a heavily glycosylated 34k protein produced in the fetal liver
cytes.
C:Genetics:
A:Gene: epo
C:Superfamily: erythropoietin
C:Keywords: glycoprotein; kidney

Query Match	Similarity	Score	Length
Best Local	81.38%	Prod. No. 1.9e-58	195
Matches	133	Conservative	12
		Mismatches	18
		Indels	1
		Gaps	1

RESULT 9
 146578
 erythropoietin - pig (fragment)
 C:/Species: Sus scrofa domestica (domestic pig)
 C:/Date: 21-Feb-1997 #sequence-revision 21-Feb-1997 #text-change 16-Jul-1999
 C:/Accession: I46578
 R:/Wen, D.; Boissel, J.
 Blood 82, 1507-1516, 1993
 A:/Title: Erythropoietin structure-function relationships: High degree of sequence homoid
 A:/Reference number: I46083; PMID:93372347; PMID:8364201
 A:/Accession: I46578
 A:/status: preliminary; translated from GB/EMBL/DBD
 A:/Molecule type: mRNA
 A:/Residues: 1-190 <WEN>
 A:/Cross-references: GB:LI0607; NID:q164445; PIDN:AAA31029.1; PID:q164446
 A:/Superfamily: erythropoietin

Query Match	80.1%;	Score 678;	DB 2;	Length 190;
Best Local Similarity	82.0%;	Pred. No. 3.3e-58;		
Matches 137; Conservative	7;	Mismatches 21;	Indels 2;	Gaps 1;

QY	1	APPICDSRYLEYLELEAKAEVAENTTGCAHCSLNEITITPDRKVNRYAKRMEYGOA	60
Db	23	APPLICDSRYLEYLELEAKAEVAENTTGCAHCSLNEITITPDRKVNRYAKRMEYGOA	82
QY	61	VEVWQGLALSEAVLRGALLVNSQWPEPLQIHDKRVSGLSRITTLRLRGAOKEAIS	120
Db	83	MEVWQGLALSEAVLRGALLVNSQWPEPLQIHDKRVSGLSRITTLRLRGAOKEAIS	142
QY	121	PPDA--ASAAPLRTTTADTFERKLEFRVYSNFLRGKILKLYTGACRTGD	165
Db	143	LPDASPSSATPLRTFAVDTLCKLFRNYSNFLRGKILKLYTGACRRD	189

RESULT 10
I46199
erythropoietin - dog (fragment)
C:Species: Canis lupus familiaris (dog)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 23-Jul-1999
C:Accession: I46199
R:Men, D.; Bolssel, J.
Blood 82, 1507-1516, 1993
A:Title: Erythropoietin structure-function relationships: High degree of sequence hom
A:Reference number: I46083; MUID:93372347; PMID:8364201
A:Accession: I46199
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-175 <MEN>
A:Cross-references: GB:I13027; NID:g290087; PIDN:AAA30842.1; PID:g552347
A:Superfamily: erythropoietin

Query Match		75.48%	Score 638:	DB 2:	Length 175:
Best Local Similarity		81.08%	Pred. No.	2.2e-54:	
Matches 124:	Conservative 13:	Mismatches 16:	Indels 0:	Gaps 0:	
QY	1	AAPPLIDDSRYLENYLLAEKFAENITTCGACHECSLNENITYPDTKVNFYAMKRREVGQA	60		
Dd	23	APPPLIDDSRYLEYRIIEARFAENVNTGGCAOGCFSFNITYPDTKVNFYMKRRDDVQA	82		
QY	61	VEWQGLALISEAVLRGOALLVNSQPWEPLQLHVDKVAQSLSLTTLRALGAQKEAIS	120		
Dd	83	LEWQGLALISEALRGOALLANASQSEFPDLHVDKVAVSLSRLSTLIRALGAQKEAMS	142		
QY	121	PPDAASAAPLRTTTADTFEKKLFERYYSNFLRGKL	153		
Dd	143	LPEASAPALRTFETVDPLCKLFRIYSNFLRGKL	175		

```

RESULT 11
G02729
Thrombopoietin - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #ext_change 05-Nov-1999
C:Accession: G02729
R:Im, S.
submitted to the EMBL Data Library, May 1996
A:Reference number: H01637
A:Accession: G02729
A:status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-353 <max>
A:Cross-references: EMBL:U59493; NID:g1401245; PIDN:AAB03392.1; PID:g1401246
C:Genetics:
C:Gene: hmpo

```

Query Match	10.68;	Score 90;	DB 2;	Length 353;
Best Local Similarity	26.3%;	Pred. No. 0.6;		
Matches 41;	Conservative 20;	Mismatches 75;	Indels 20;	Gaps 5;
1 APRRLICDSRVLRLRYLLAEAKENAIITTCGAEHCISLLENITVPTTKVNFYAMKRMEVGOQA 60				


```
Db 24 APP--ACDLRVLSKLRSDSHVLSHSKLSOCPEVHPRLPTFVLLPAVDFSGEWTQMEETKA 81
QY 61 VEWMOGLALISEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGLSITLTLRALGAKXA 118
Db 82 QDILGAVTLLEGVMAARQGLGPTCLSSLLGQLSGVRLRLGALQSL-----LGTQ--- 132
QY 119 ISPPDASAPLRITTTADTFKRLFRVYSNFLRGKLL 154
Db 133 -LPPQG-----RTTAHKDPNAIFLSFQHLRLGKVR 161

RESULT 12
180105
thrombopoietin precursor - human
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 20-Jun-2000
C:Accession: I59281; I80105; S45331; S48740; I38672; I52610
R:Poster, D.C.; Sprecher, C.A.; Grant, F.J.; Kramer, J.M.; Kujper, J.L.; Holly, R.D.; W
Proc. Natl. Acad. Sci. U.S.A. 91, 13023-13027, 1994
A:Title: Human thrombopoietin: gene structure, cDNA sequence, expression, and chromosom
A:Reference number: I59281; MUID:95108091; PMID:7809166
A:Accession: I59281
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-353 <RES>
A:Cross-references: GB:I36051; NID:9533214; PIDN:AAC37568.1; PID:9533215
A:Accession: I80105
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-353 <RES>
A:Cross-references: GB:I36052; NID:9533216; PIDN:AAC37566.1; PID:9533217
R:de Sauvage, F.J.; Hoss, P.E.; Spencer, S.D.; Melloy, B.E.; Gurney, A.L.; Spencer, S.A.
D.V.; Eaton, D.L.
Nature 369, 533-538, 1994
A:Title: Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl ligand.
A:Reference number: S45331; MUID:94261202; PMID:8202154
A:Accession: S45331
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-353 <SAU>
A:Cross-references: GB:I33410; NID:9506826; PIDN:AAA9857.1; PID:9506827
R:Solmajer, Y.; Akahori, H.; Seki, N.; Horl, T.; Ogami, K.; Kato, T.; Shimada, Y.; Kawamura
FBS Lett. 353, 57-61, 1994
A:Title: Molecular cloning and chromosomal localization of the human thrombopoietin gene
A:Reference number: S48740; MUID:95010765; PMID:7926023
A:Accession: S48740
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <SOH>
A:Cross-references: GB:D32046; NID:9577319; PIDN:BA06807.1; PID:9577320
R:Barclay, T.D.; Bogenberger, J.; Hunt, P.; Li, Y.S.; Lu, H.S.; Martin, F.; Chang, M.S.;
Cell 77, 1117-1124, 1994
A:Title: Identification and cloning of a megakaryocyte growth and development factor tha
A:Reference number: A54463; MUID:94291201; PMID:8020099
A:Accession: I38672
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-112, 'E', 114-353 <RE3>
A:Cross-references: EMBL:U11025; NID:9511223; PIDN:AAA5055.1; PID:9558078
R:Gurney, A.L.; Huang, W.J.; Xie, M.H.; Melloy, B.E.; Eaton, D.L.; de Sauvage, F.J.
Blood 85, 981-988, 1995
A:Title: Genomic structure, chromosomal localization, and conserved alternative splice f
A:Reference number: I52610; MUID:95152076; PMID:7845319
A:Accession: I52610
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-353 <REA>
A:Cross-references: GB:S76771; NID:9914225; PIDN:AAB33390.1; PID:9914226
C:Genetics:
A:Gene: GDB:THPO; MGD
A:Cross-references: GDB:374007; OMIM:600044
```

```
A:Map position: 3q26.3-3q27
A:introns: 5/1; 47/3; 76/3; 132/3
C:Keywords: alternative splicing; cytokine; glycoprotein
Query Match 10.5%; Score 89; DB 2; Length 353;
Best Local Similarity 26.3%; Pred. No. 0.75;
Matches 41; Conservative 20; Mismatches 75; Indels 20; Gaps 5;

QY 1 APPRLICSRVLEKRLLEKAEENITTCGAECISNENITVPDTKYNFAKRMVEVGOA 60
Db 24 APP--ACDLRVLSKLRSDSHVLSHSKLSOCPEVHPRLPTFVLLPAVDFSGEWTQMEETKA 81
QY 61 VEWMOGLALISEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGLSITLTLRALGAKXA 118
Db 82 QDILGAVTLLEGVMAARQGLGPTCLSSLLGQLSGVRLRLGALQSL-----LGTQ--- 132
QY 119 ISPPDASAPLRITTTADTFKRLFRVYSNFLRGKLL 154
Db 133 -LPPQG-----RTTAHKDPNAIFLSFQHLRLGKVR 161

RESULT 13
AB0323
ribonucleoside-diphosphate reductase (EC 1.17.4.1) beta chain [Imported] - Yersinia p
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C:Accession: AB0323
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tiltball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0323
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-323 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92889.1; PID:915980631; GSPDB:GN00175
C:Genetics:
A:Gene: nrdf
C:Superfamily: ribonucleoside-diphosphate reductase beta
C:Keywords: oxidoreductase

Query Match 10.4%; Score 88; DB 2; Length 323;
Best Local Similarity 25.2%; Pred. No. 0.84;
Matches 34; Conservative 20; Mismatches 59; Indels 22; Gaps 5;

QY 38 NITVPDTKYNFAKRMVEVGOA VEWMOGLALISEAVL RGQALLVNSSQPWEPLQLHVD- 96
Db 2 NVYKPTRISALNMWKIE-DDKDLEVMN--RLTSNFWLPEKYPKPLSNDIFSMTLTPHEQQ 58
QY 97 ---KAVSGLSITLTLRALGAA--KEAISPPDASAPLRITTTADTFKRLFRVYSNFLR 150
Db 59 LTRIVETGTLTDITONTGLAPALIKDITPHEAIFSNISFEAVHARSYSIFSTL-- 116
QY 151 GKILKLYTGACTGCD 165
Db 117 -----CLTSD 121

RESULT 14
AE0959
Solute binding receptor protein [Imported] - Salmonella enterica subsp. enterica sero
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AE0959
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
S.; Moulle, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
```

A:Reference number: AB0502; PMID:11677608
A:Accession: AE0959
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1346 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD03169.1; PID:g16504804; GSPDB:GN00176
C:Genetics: STY3952

Query Match 10.3%; Score 87.5; DB 2; Length 346;
Best Local Similarity 26.7%; Pred. No. 1;
Matches 44; Conservative 22; Mismatches 48; Indels 51; Gaps 9;
QY 10 RYLERYLEKAEKNITTG--CAEHCISNE--NITVPDTRKVNFTYAMKRMEVGOAVEVWQ 65
DB 217 RNLQEMLEHPRDANVAGSAIAEAAGRGMDITPLTIVSTL-----THQVXR 267
QY 66 GLALISEAVLRGQALLVNSQ-PMEPLQLHVDKAVSGLSLTLLRALGQ--KEAISP 122
DB 268 GLK-----RGHILMALSQAMQ-----GELAITOSTKYVLOGQVPEWISPP 309
QY 123 -----DASAAPLRTITADPFRKLPVYSNFRGKRLTYGEA 160
DB 310 VLITLNNADSARVRSISPPGFRPY-----LYQITSEA 344

RESULT 15
A55530
megakaryocyte growth and development factor, long form - human
N:Alternate names: MPL ligand, long form
C:Species: Homo sapiens (man)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 07-May-1999
C:Accession: A55530
R:Chang, M.; McIninch, J.; Basu, R.; Shutter, J.; Hsu, R.; Perkins, C.; Mar, V.; Suggs, S.
J. Biol. Chem. 270, 511-514, 1995
A:Title: Cloning and characterization of the human megakaryocyte growth and development
A:Reference number: A55530; MUID:95122483; PMID:7822271
A:Accession: A55530
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-286 <CHA>
A:Cross-references: GB:U17071
C:Genetics:
A:Gene: MGDF
A:Map position: 3q26.3
C:Keywords: alternative splicing; cytokine

Query Match 10.2%; Score 86; DB 2; Length 286;
Best Local Similarity 26.6%; Pred. No. 1.1;
Matches 41; Conservative 18; Mismatches 75; Indels 20; Gaps 5;
QY 1 APPRLICDSRVLEKYLEKAEKNITTGCAEHCISNENITVPDTRKVNFTYAMKRMEVGOA 60
DB 24 APP--ACDLRVLSKLLRDSHVLSRSLSQCEVHRLPTVLPAVDPSLGEWKTQMEETKA 81
QY 61 VEWMOGLALISEAVL--RGQALLVNSQPMELDLHVDKAVSGLSLTLLRALGAKKA 118
DB 82 ODILGAVTLLEGVMAARGQLGPTCLSSLLGQLSGQVRLLLGALQSL-----LQTQ--- 132
QY 119 ISPPDASAAPLRTITADPFRKLPVYSNFRGK 152
DB 133 -LPFG-----RTTAHKDPNAIFLSFQHLLRGK 159

Search completed: January 2, 2003, 15:13:26
Job time : 14.9577 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:07:09 : Search time 8.47432 Seconds

(without alignments)
807.569 Million cell updates/sec

Title: US-09-853-731-1

Perfect score: 846
Sequence: 1 APPRLICDSRVLEERYLEAK.....SNFLGKLIKITYGACRTGD 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	846	100.0	193	EPO_HUMAN	P01588 homo sapien
2	764.5	90.4	192	EPO_MACFA	P07865 macaca fasc
3	759.5	89.8	192	EPO_MACMU	O28513 macaca mula
4	706	83.5	192	EPO_FETCA	P33708 felis silve
5	701	82.9	192	EPO_RAT	P29676 rattus norv
6	692.5	81.9	192	EPO_BOVIN	P48617 bos taurus
7	689	81.4	192	EPO_MOUSE	P07321 mus musculu
8	685.5	81.0	194	EPO_SHEEP	P33709 ovis aries
9	678	80.1	190	EPO_PIG	P49157 sus scrofa
10	638	75.4	175	EPO_CANFA	P33707 canis famill
11	109	12.9	352	TPO_CANFA	P42705 canis famill
12	89	10.5	353	TPO_HUMAN	P40225 homo sapien
13	80.5	9.5	3033	POLG_HCVJ8	P26661 h genome po
14	78.5	9.3	897	EPL5_MOUSE	P42667 mus musculu
15	78	9.2	747	ATCS_SYNP7	P37279 synecococc
16	77	9.1	548	CH60_BUCAP	O51832 buchnera ap
17	76.5	9.0	386	CTBP_DROME	O46036 drosophila
18	76	9.0	551	CH60_BUCAP	O59177 buchnera ap
19	74.5	8.8	552	CH60_PSEST	O33500 pseudomonas
20	74.5	8.8	809	HRPB_ECOLI	P37324 escherichia
21	74	8.7	326	TPO_RAT	P49745 rattus norv
22	74	8.7	1564	PDR4_YEAST	P51533 saccharomyc
23	73.5	8.7	830	GC12_MOUSE	O91nd2 mus musculu
24	73.5	8.7	837	GC12_HUMAN	O92nd3 homo sapien
25	73	8.6	263	YH25_DEIRA	O83030 delnocoocys
26	73	8.6	1089	IMB3_YEAST	P32337 saccharomyc
27	72.5	8.6	756	SLAP_ACEKI	P22258 acetogenium
28	72	8.5	362	TPO_MOUSE	P40226 mus musculu
29	72	8.5	548	CH60_BUCAI	P25750 buchnera ap
30	70.5	8.3	217	MYND_ECOLI	P76909 escherichia
31	70.5	8.3	381	MODD_MYCAV	O48919 mycobacteri
32	70.5	8.3	543	CH60_BARBA	P35635 bartonella
33	70.5	8.3	896	EPL5_HUMAN	P42566 homo sapien

ALIGNMENTS

RESULT 1	ID	EPO_HUMAN	STANDARD:	PRT:	193 AA.
AC	P01588; Q9UHA0; Q9UE25; Q9UD20;				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	21-JUL-1986 (Rel. 01, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Erythropoietin precursor (Epoetin).				
GN	EPO.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_Taxid-9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-85137899; PubMed-3838366;				
RA	Jacobs K., Shoemaker C., Ruderstorf R., Neill S.D., Kaufman R.J.,				
RA	Mufson A., Seehra J., Jones S.S., Hewick R., Fritsch E.F.,				
RA	Kawakita M., Shimizu T., Miyake T.;				
RT	"Isolation and characterization of genomic and cDNA clones of human				
RT	erythropoietin.";				
RL	Nature 313:806-810(1985).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-86067948; PubMed-3865178;				
RA	Lin F.-K., Suggs S., Lin C.-H., Browne J.K., Egrle J.C.,				
RA	Chen K.K., Fox G.M., Martin F., Stabinsky Z., Bedrawi S.M., Lai P.-H.,				
RA	Goldwasser E.;				
RT	"Cloning and expression of the human erythropoietin gene.";				
RT	Proc. Natl. Acad. Sci. U.S.A. 82:7580-7584(1985).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-99018118; PubMed-9799793;				
RA	Gloekner G., Scherer S., Schattevoy R., Borlight A., Weber J.,				
RA	Tsui L.-C., Rosenthal A.;				
RT	"Large-scale sequencing of two regions in human chromosome 7q22:				
RT	analysis of 650 kb of genomic sequence around the EPO and CDT1 loci				
RT	reveals 17 genes.";				
RL	Genome Res. 8:1060-1073(1998).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Rupert J.L., Hochachka P.W.;				
RT	"Erythropoietin gene sequence in the Quechua, a high altitude native				
RT	population.";				
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE OF 58-193 FROM N.A., AND VARIANTS HEPATOCELLULAR CARCINOMA.				
RX	MEDLINE-93384593; PubMed-8396923;				
RA	Funakoshi A., Muta H., Baba T., Shimizu S.;				
RT	"Gene expression of mutant erythropoietin in hepatocellular				
RT	carcinoma.";				
RL	Biochem. Biophys. Res. Commun. 195:717-722(1993).				
RN	[6]				
RP	SEQUENCE OF 28-193, AND DISULFIDE BONDS.				
RT	TISSUE-Urine;				
RC	MEDLINE-86140080; PubMed-3949763;				

34	70	8.3	319	1	RIR4_SALTY	P17424 salmonella
35	70	8.3	547	1	CH60_LEGPN	P26878 legionella
36	69.5	8.2	907	1	GACS_PSEST	P48027 pseudomonas
37	69.5	8.2	1327	1	TNR1_HUMAN	O95271 homo sapien
38	69	8.2	353	1	NADA_YERPE	O82978 yeastinia pe
39	69	8.2	544	1	CH60_FRATU	P94798 francisella
40	69	8.2	552	1	CH60_COXBU	P19421 coxiella bu
41	69	8.2	1128	1	DNRT_HSVSA	P24910 herpesvirus
42	68.5	8.1	360	1	CYS2_MAIZE	O10717 zea mays (m
43	68.5	8.1	418	1	GUN2_TIRRE	P07982 trichoderma
44	68.5	8.1	778	1	RG12_MOUSE	O61193 mus musculu
45	68	8.0	366	1	QVEA_SYNY3	O58850 synecocyst

RA Lai P.H., Everett R., Wang F.F., Arakawa T., Goldwasser E.;
 "Structural characterization of human erythropoietin.";
 J Biol. Chem. 261:3116-3121(1986).
 RL
 RP PRELIMINARY SEQUENCE OF 28-57.
 RX MEDLINE-6413751; PubMed-6698989;
 RA Tanigawa S., Hirade K., Onoda H., Sasaki R., Chiba H., Ueda M.,
 Soto M.;
 "Isolation of human erythropoietin with monoclonal antibodies.";
 J Biol. Chem. 259:2707-2710(1984).
 RL
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE-86133657; PubMed-346214;
 RA Takeuchi M., Takasaki S., Miyazaki H., Kato T., Hoshi S., Kochibe N.,
 Kobata A.;
 "Comparative study of the asparagine-linked sugar chains of human
 erythropoietins purified from urine and the culture medium of human
 recombinant Chinese hamster ovary cells.";
 J Biol. Chem. 263:3657-3663(1988).
 RL
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE-89118279; PubMed-3219367;
 RA Sasaki H., Ochi N., Dell A., Fukuda M.;
 "Site-specific glycosylation of human recombinant erythropoietin:
 analysis of glycopeptides or peptides at each glycosylation site by
 fast atom bombardment mass spectrometry.";
 J Biochemistry 27:8618-8626(1988).
 RL
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE-92314463; PubMed-1820196;
 RA Takeuchi M., Kobata A.;
 "Erythropoietin and functional roles of the sugar chains of human
 erythropoietins.";
 J Biol. Chem. 263:3657-3663(1988).
 RL
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE-98445092; PubMed-9774108;
 RA Shen R.S., Reid S.W., Li C., Cheetham J.C., Aoki K.H., Liu B.,
 Zhan H., Osslund T.D., Chiaro A.J., Zhang J., Piner-Moore J.,
 Elliott S., Stoney K., Katz B.A., Matthews D.J., Wendoloski J.J.,
 Fridge J., Stroud R.M.;
 "Efficiency of signalling through cytokine receptors depends
 critically on receptor orientation.";
 Nature 393:511-516(1998).
 RL
 RP REGULATION OF ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
 PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
 CC
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC
 CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
 CC
 CC -1- AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
 CC
 CC PHARMACEUTICAL: Used for the treatment of anemia. Available under
 CC the names Epogen (Amgen), Epogin (Chugai), Epomax (Eli Lilly), and Procrit
 CC (Janssen-Cilag). Variations in the glycosylation pattern of Epo
 CC distinguishes these products. Epogen, Epogin, Epex and Procrit
 CC are generically known as epoetin alfa, NeoRecormon and Recormon as
 CC epoetin beta and Epomax as epoetin omega.
 CC
 CC -1- SIMILARITY: BELONGS TO THE Epo / TPO FAMILY.
 CC
 CC -1- DATABASE: NAME-RE: SwissProt, cytochrome source book: EPO;
 CC WWW="http://www.indsystems.com/asp/g_sitbuilder.asp?bodyid=197".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: X02158; CAA26095.1; -
 CC EMBL: X02157; CAA26094.1; -
 CC EMBL: M1319; AAA52400.1; -
 CC EMBL: AF053356; AAC78791.1; -
 CC

DR EMBL: AF202308; AAF23132.1; -
 DR EMBL: AF202306; AAF23132.1; JOINED.
 DR EMBL: AF202307; AAF23132.1; JOINED.
 DR EMBL: AF202310; AAF23133.1; -
 DR EMBL: AF202309; AAF23133.1; JOINED.
 DR EMBL: AF202311; AAF17572.1; -
 DR EMBL: AF202314; AAF23134.1; -
 DR EMBL: AF202312; AAF23134.1; JOINED.
 DR EMBL: AF202313; AAF23134.1; JOINED.
 DR EMBL: S65458; AAD13964.1; -
 DR PIR: A01855; ZUH0.
 DR PIR: A25384; A25384.
 DR PIR: A24744; A24744.
 DR PIR: A22210; A22210.
 DR PDB: 1EER; 01-OCT-99.
 DR GlycoSuiteDB: P01588; -
 DR GeneW: HGNC:3415; EPO.
 DR MIM: 133170; -
 DR InterPro: IPR001323; EPO.TPO.
 DR InterPro: IPR003013; Erythropo.
 DR Pfam: PF00758; EPO.TPO.1.
 DR PRINTS: PR00272; ERYTHROPIN.
 DR PROSITE: PS00817; EPO.TPO.1.
 DR Erythrocyte maturation; Glycoprotein; Hormone; Signal; Pharmaceutical;
 DR 3D-structure.
 DR SIGNAL 1 27
 DR CHAIN 28 193
 DR PROPE 190 193
 DR DISULFID 34 188
 DR DISULFID 56 60
 DR CARBOHYD 51 51
 DR CARBOHYD 65 65
 DR CARBOHYD 110 110
 DR CARBOHYD 153 153
 DR CARBOHYD 131 132
 DR VARIANT 149 149
 DR VARIANT 149 149
 DR CONFLICT 40 40
 DR CONFLICT 85 85
 DR CONFLICT 140 140
 DR SEQUENCE 193 AA; 21306 MW; C91F0B4C26A52033 CRC64;
 Query Match 100.0%; Score 846; DB 1; Length 193;
 Best Local Similarity 100.0%; Pred No. 1.9e-76; Mismatches 0; Gaps 0;
 Matches 165; Conservative 0; Indels 0;
 Oy 1 APRILICRYLEYLELLEAKAEVITTCAGHCSINENIVPTKVFYFMKREVGQQA 60
 Db 28 APRILICRYLEYLELLEAKAEVITTCAGHCSINENIVPTKVFYFMKREVGQQA 87
 Oy 61 VEWVQGLALSEAVLRQALLVNSQWPEPLQHVKAAYGSLNLTLLPALCAQKREAS 120
 Db 88 VEWVQGLALSEAVLRQALLVNSQWPEPLQHVKAAYGSLNLTLLPALCAQKREAS 147
 Oy 121 PPDAASAPLFTITADPFRKLFVSNFLGKLTLYGECGRGD 165
 Db 148 PPDAASAPLFTITADPFRKLFVSNFLGKLTLYGECGRGD 192
 RESULT 2
 EPO_MACFA STANDARD; PRT: 192 AA.
 AC P07865;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Erythropoietin precursor.

```

GN EPO.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8705236; PubMed=2877922;
RA Lin E.-K., Lin C.-H., Lai P.-H., Browne J.K., Egrie J.C., Smalling R.,
RA Fox G.M., Chen K.K., Castro M., Suggs S.;
RT "Monkey erythropoietin gene: cloning, expression and comparison with
RT the human erythropoietin gene."
RL Gene 44:201-209(1986).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL; M18189; AAA36841.1; -.
DR PIR; J00173; J00173.
DR HSSP; P01588; ICN4.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythropo.
DR Pfam; PF00758; EPO_TPO.1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO.1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 27
FT CHAIN 28 192
FT DISULFID 34 187
FT CARBOHYD 51 51
FT CARBOHYD 65 65
FT CARBOHYD 110 110
FT CARBOHYD 152 152
SQ SEQUENCE 192 AA; 21113 MW; E8A900F442AD4522 CRC64;

Query Match 90.4%; Score 764.5; DB 1; Length 192;
Best Local Similarity 91.5%; Pred. No. 2,2e+68;
Matches 151; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 1 APRRLICDSRVLEERYLLLEKAEENITTCGAHCNSLNIENITVPTKYNFAFKRMVEGGOA 60
DB 28 APRRLICDSRVLEERYLLLEKAEENITTCGAHCNSLNIENITVPTKYNFAFKRMVEGGOA 87
QY 61 VEVWQGLALISEAVLRGQALVNSSQPEPQLQHVDAKAVSGRLSTLTLLRALGAQKEALS 120
DB 88 VEVWQGLALISEAVLRGQAVLVNSSQPEPQLQHMDKALSGRLSTLTLLRALGAQ- EALS 146
QY 121 PDDASAPLPRTITADTFCKLFRVYSNPLRGKILKLTGTGACRTGD 165
DB 147 LPDASAPLPRTITADTFCKLFRVYSNPLRGKILKLTGTGACRRGD 191

RESULT 3
EPO_MACMU
ID EPO_MACMU STANDARD; PRT; 192 AA.
AC 028513;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE Erythropoietin precursor.
GN EPO.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93372347; PubMed=8364201;
RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA Czelusniak J., Goodman M., Bunn H.F.;
RT "Erythropoietin structure-function relationships: high degree of
RT sequence homology among mammals."
RL Blood 82:1507-1516(1993).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL; L10609; AAA36842.1; -.
DR HSSP; P01588; ICN4.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythropo.
DR Pfam; PF00758; EPO_TPO.1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO.1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 27
FT CHAIN 28 192
FT DISULFID 34 187
FT CARBOHYD 51 51
FT CARBOHYD 65 65
FT CARBOHYD 110 110
FT CARBOHYD 152 152
SQ SEQUENCE 192 AA; 21081 MW; 275560A264628CD1 CRC64;

Query Match 89.8%; Score 759.5; DB 1; Length 192;
Best Local Similarity 90.3%; Pred. No. 6,7e+68;
Matches 149; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

QY 1 APRRLICDSRVLEERYLLLEKAEENITTCGAHCNSLNIENITVPTKYNFAFKRMVEGGOA 60
DB 28 APRRLICDSRVLEERYLLLEKAEENITTCGAHCNSLNIENITVPTKYNFAFKRMVEGGOA 87
QY 61 VEVWQGLALISEAVLRGQALVNSSQPEPQLQHVDAKAVSGRLSTLTLLRALGAQKEALS 120
DB 88 VEVWQGLALISEAVLRGQAVLVNSSQPEPQLQHMDKALSGRLSTLTLLRALGAQ- EALS 146
QY 121 PDDASAPLPRTITADTFCKLFRVYSNPLRGKILKLTGTGACRTGD 165
DB 147 LPDASAPLPRTITADTFCKLFRVYSNPLRGKILKLTGTGACRRGD 191

RESULT 4
EPO_FELCA
ID EPO_FELCA STANDARD; PRT; 192 AA.
AC P33708;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)

```

DE 16-OCT-2001 (Rel. 40, last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Goodman R.E., Bell R.G.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 5-192 FROM N.A.
RX MEDLINE=93372347; PubMed=8364201;
RA Men D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA Czelusniak J., Goodman M., Bunn H.F.;
RT *Erythropoietin structure-function relationships: high degree of
RT sequence homology among mammals.";
RL Blood 82:1507-1516(1993).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)

DR EMBL; U00685; AAA18282.1; -;
DR EMBL; L10606; AAA30807.1; -;
DR HSSP; P01588; 1CN4.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythropo.
DR Pfam; PF00758; EPO_TPO.1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO.1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 26
FT CHAIN 27 192
FT DISULFID 33 187
FT DISULFID 55 187
FT CARBOHYD 50 59
FT CARBOHYD 64 59
FT CARBOHYD 109 109
FT CONFILCT 44 109
SQ SEQUENCE 192 AA; 20914 MW; 61C5EA0FE937293 CRC64;
Query Match 83.5%; Score 706; DB 1; Length 192;
Best Local Similarity 83.6%; Pred. No. 1,36-62;
Matches 138; Conservatave 9; Mismatches 18; Indels 0; Gaps 0;

EPO_RAT
ID EPO_RAT STANDARD; PRT; 192 AA.
RT 29676; P70504;
DR 01-APR-1993 (Rel. 25, Created)
DR 01-APR-1993 (Rel. 25, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Missar; TISSUE=Kidney;
RX MEDLINE=93042015; PubMed=1420369;
RA Nagao M., Suga H., Okano M., Masuda S., Narita H., Ikura K.,
RA Sasaki R.;
RT Nucleotide sequence of rat erythropoietin.";
RL Biochim. Biophys. Acta 1171:99-102(1992).
RN [2]
RP SEQUENCE OF 4-192 FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=93372347; PubMed=8364201;
RA Men D., Boissel J.P.R., Tracy T.E., Mulcahy L.S., Czelusniak J.,
RA Goodman M., Bunn H.F.;
RT *Erythropoietin structure-function relationships: high degree of
RT sequence homology among mammals.";
RL Blood 82:1507-1516(1993).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)

DR EMBL; D10763; BAA01593.1; -;
DR EMBL; L10608; BAA41126.1; -;
DR PIR; S28148; S28148.
DR HSSP; P01586; 1CN4.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythropo.
DR Pfam; PF00758; EPO_TPO.1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO.1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 26
FT CHAIN 27 192
FT DISULFID 33 187
FT DISULFID 50 59
FT CARBOHYD 64 59
FT CARBOHYD 109 109
SQ SEQUENCE 192 AA; 21286 MW; 3EA032737E7D2443 CRC64;
Query Match 82.9%; Score 701; DB 1; Length 192;
Best Local Similarity 82.4%; Pred. No. 4e-62;
Matches 136; Conservatave 13; Mismatches 16; Indels 0; Gaps 0;

[illegible]

```

Db      86 LEWGGGLDLSFALLRGQLLNANASQPCBALHVDKAVSGLRSLTSLRALGAKKAITS 145
Oy      121 PDDA-SAAPLRTITADTFERKLFERYYSNFLRCKLTYTGACRTGD 165
Db      146 LPDAPPSAAPLRAFTVDALSKFLFRITYSNRLRGKLTLYTGACRRGD 191

RESULT 7
EPO_MOUSE
ID EPO_MOUSE STANDARD: PRT; 192 AA.
AC P07321.
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87039105; PubMed=3773894;
RA Shoemaker C.B., Mitscock L.D.;
RT "Murine erythropoietin gene: cloning, expression, and human gene
RT homology";
RL Mol. Cell. Biol. 6:849-858(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87039104; PubMed=3022133;
RA McDonald J.D., Lin F.-K., Goldwasser E.;
RT "Cloning, sequencing, and evolutionary analysis of the mouse
RT erythropoietin gene";
RL Mol. Cell. Biol. 6:842-848(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=129/Sv;
RC MEDLINE=21138439; PubMed=11239002;
RA Wilson M.D., Riemer C., Martindale D.W., Schnupf P., Boright A.P.,
RA Cheng T.L., Hardy D.M., Schwartz S., Scherer S.W., Tsui L.-C.,
RA Miller W., Koop B.F.;
RT "Comparative analysis of the gene-dense ACHE/TR2 region on human
RT chromosome 7q22 with the orthologous region on mouse chromosome 5.";
RL Nucleic Acids Res. 29:1352-1365(2001).
RN [4]
RP SEQUENCE OF 1-52 FROM N.A.
RC STRAIN=ICFW;
RX MEDLINE=96030528; PubMed=9365246;
RA Chretien S., Duprez V., Maouche L., Gisselbrecht S., Mayeux P.,
RA Lacombe C.;
RT "Abnormal erythropoietin (Epo) gene expression in the murine
RT erythroleukemia IM32 cells results from a rearrangement between the
RT G-protein beta2 subunit gene and the Epo gene.";
RL Oncogene 15:1995-1999(1997).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/,
CC or send an email to license@sib-sib.ch)
CC -----
DR EMBL; M12482; AAA37568.1; -.
DR EMBL; M12930; AAA37570.1; -.

```

DR	EMBL:	AF312033;	AARK8825.1;	-
DR	EMBL:	Y11971;	CAA72207.1;	-
DR	PIR:	A24901;	A24901.	
DR	PIR:	A24902;	A24902.	
DR	HSSP:	P01588;	ICM4.	
DR	MCD:	MGI:95407;	Epo.	
DR	InterPro:	IIPR01323;	EPO_TPO.	
DR	InterPro:	IIPR003013;	Erythropun.	
DR	Pfam:	PF00758;	EPO_TPO.1	
DR	PRINTS:	PR00722;	ERYTHROPTN.	
DR	PROSITE:	PS00817;	EPO_TPO.1	
KW	Erythrocyte maturation; Glycoprotein; Hormone; Signal.			
FT	SIGNAL	1	26	
FT	CHAIN	27	192	
FT	DISULFID	33	187	ERYTHROPOIETIN.
FT	CARBOHYD	50	50	BY SIMILARITY.
FT	CARBOHYD	64	64	N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT	CARBOHYD	109	109	N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FO	SEQUENCE	192 AA;	21365 MW;	65F94E214ED0DEFECRC64;

[illegible]

CC	RT	RA	RC	RP	RL	RN	RESULT 8
CC	RT	RA	RC	RP	RL	RN	EPO_SHEEP
CC	RT	RA	RC	RP	RL	RN	ID EPO_SHEEP
CC	RT	RA	RC	RP	RL	RN	P33709; 028572; STANDARD; PRT; 194 AA.
CC	RT	RA	RC	RP	RL	RN	01-FEB-1994 (Rel. 28, Created)
CC	RT	RA	RC	RP	RL	RN	01-FEB-1994 (Rel. 28, Last sequence update)
CC	RT	RA	RC	RP	RL	RN	16-OCT-2001 (Rel. 40, Last annotation update)
CC	RT	RA	RC	RP	RL	RN	Erythropoietin precursor.
CC	RT	RA	RC	RP	RL	RN	EPO.
CC	RT	RA	RC	RP	RL	RN	Ovis aries (Sheep).
CC	RT	RA	RC	RP	RL	RN	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	RT	RA	RC	RP	RL	RN	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
CC	RT	RA	RC	RP	RL	RN	Bovidae; Caprinae; Ovis.
CC	RT	RA	RC	RP	RL	RN	NCBI_TaxID=9940;
CC	RT	RA	RC	RP	RL	RN	[1]
CC	RT	RA	RC	RP	RL	RN	SEQUENCE FROM N.A.
CC	RT	RA	RC	RP	RL	RN	TISSUE=Kidney;
CC	RT	RA	RC	RP	RL	RN	MEDLINE=93351736; PubMed=8349021;
CC	RT	RA	RC	RP	RL	RN	Fu P., Evans B., Lim G.B., Moritz K., Wintour M.E.;
CC	RT	RA	RC	RP	RL	RN	"The sheep erythropoietin gene: molecular cloning and effect of
CC	RT	RA	RC	RP	RL	RN	hemorrhage on plasma erythropoietin and renal/liver messenger RNA in
CC	RT	RA	RC	RP	RL	RN	adult sheep.";
CC	RT	RA	RC	RP	RL	RN	[2]
CC	RT	RA	RC	RP	RL	RN	Mol. Cell. Endocrinol. 93:107-116(1993).
CC	RT	RA	RC	RP	RL	RN	SEQUENCE OF 4-194 FROM N.A.
CC	RT	RA	RC	RP	RL	RN	TISSUE=Kidney;
CC	RT	RA	RC	RP	RL	RN	MEDLINE=93372347; PubMed=8364201;
CC	RT	RA	RC	RP	RL	RN	Wen D., Bolissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
CC	RT	RA	RC	RP	RL	RN	Czelusnak J., Goodman M., Bunn H.F.;
CC	RT	RA	RC	RP	RL	RN	"Erythropoietin structure-function relationships: high degree of
CC	RT	RA	RC	RP	RL	RN	sequence homology among mammals.";
CC	RT	RA	RC	RP	RL	RN	Blood 82:1507-1516(1993).
CC	RT	RA	RC	RP	RL	RN	-1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC	RT	RA	RC	RP	RL	RN	REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC	RT	RA	RC	RP	RL	RN	PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC	RT	RA	RC	RP	RL	RN	-1- SUBCELLULAR LOCATION: secreted.

CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on list-
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR	EMBL; Z24681; CAA80848.1; -.
DR	EMBL; L10610; AAA31518.1; -.
DR	HSSP; P01588; ICN4.
DR	InterPro; IPR001323; EPO.TPO.
DR	InterPro; IPR003013; Erythropn.
DR	Pfam; PF00758; EPO.TPO.1
DR	PRINTS; PR00272; ERYTHROPTN.
DR	PROSITE; PS00617; EPO.TPO.1.
KW	Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT	SIGNAL
ET	CHAIN
ET	DISULFID
FT	DISULFID
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CONFLICT
FT	CONFLICT
XQ	SEQUENCE

Query Match	Similarity	Score	DB 1:	Length
Best Local	81.9%	Pred. 1.4e-60		
Matches	136; Conservative	9; Mismatches	20; Indels	1; Gaps
QY	1	APPRILCSRYLERYLLLEAKAEANTITTCGAHCISINENITVPDTKYNYFAMKRMVEYGOQA	60	
Db	28	APPRILCSRYLERYLLLEAREENNTMCACBSCSSENTVPDTKYNYFAMKRMVEYGOQA	87	
QY	61	VEWOGALLSEAVLRGALLVYSSQPEWPIQLHYDKAVSGSRITTLTIRALGAQKEATS	120	
Db	88	LEWVGALLSEAVLRGALLVYSSQPEWPIQLHYDKAVSGSRITTLTIRALGAQKEATS	147	
QY	121	PPDAA-SAAPLRITTTADPTFRKLFERYYSNFLRKLKLYTGEACRTGD	165	
Db	148	LPDPTPSAAPLRITFTVDALSKIFRYTSNFLRKLKLYTGEACRTGD	193	

RESULT 9			
EPO_PIG	STANDARD:	PRT:	190 AA.
ID	EPO_PIG		
AC	P49157;		
DT	01-FEB-1996 (Rel. 33, Created)		
DT	01-FEB-1996 (Rel. 33, last sequence update)		
DT	16-OCT-2001 (Rel. 40, last annotation update)		
DE	Erythropoietin precursor (Fragment).		
GN	EPO.		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxId=9823;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Kidney;		
RX	MEDLINE=93372347; PubMed=8364201;		
RA	Wen D., Bollstedt J.P., Tracy T.E., Gruninger R.H., Mulcahy L.S.,		
RA	Czelusniak J., Goodman M., Bunn H.F.;		
RT	"Erythropoietin structure-function relationships: high degree of		
RT	sequence homology among mammals.";		
RL	Blood 82:1507-1516 (1993).		
CC	-1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE		
CC	REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A		


```

CC      PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC      AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC      -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: L10607; AAA31029.1; -.
CC      DR      HSSP: P01588; 1CN4.
CC      DR      InterPro: IPR001323; EPO_TPO.
CC      DR      Pfam: PF00758; EPO_TPO; 1.
CC      KW      Prosite: PS00817; EPO_TPO; 1.
CC      KW      Erythrocyte maturation; Glycoprotein; Hormone; Signal.
CC      FT      NON_TER 1
CC      FT      SIGNAL <1 22
CC      FT      CHAIN 23 190
CC      FT      DISULFID 29 185
CC      FT      DISULFID 51 55
CC      FT      CARBOHYD 46 46
CC      FT      CARBOHYD 60 60
CC      FT      CARBOHYD 105 105
CC      FT      CARBOHYD 168 168
CC      SQ      SEQUENCE 190 AA; 20888 MW; A75BDCCE5077E2A CRC64;

Query Match
Best Local Similarity 80.1%; Score 678; DB 1; Length 190;
Matches 137; Conservative 7; Mismatches 21; Indels 2; Gaps 1;

OY      1 APPRLICDSRVLEERYLLLEKAEKNITTCGAEHCSLNENITVPDTKVFYAMKMEYGOQA 60
        |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB      23 APPRLICDSRVLEERYLLLEKAEKNITTCGAEHCSLNENITVPDTKVFYAMKMEYGOQA 82
OY      61 VEWOGALLSEAVLNGQALLVNSQWPEPLQHDVKAVSGLSLTLTLRALGAQKEAIS 120
        :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB      83 LEWOGALLSEAVLNGQALLVNSQWPEPLQHDVKAVSGLSLTLTLRALGAQKEAIS 142
OY      121 PPDA--ASAAPLRTTADTFKRLFRYSNPLRGKLTLYGECRCRD 165
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB      143 LPDASPSATPLRTAVDTLCKLFRNYSNPLRGKLTLYGECRCRD 189

RESULT 10
EPO_CANFA
ID      EPO_CANFA          STANDARD;          PRT;          175 AA.
AC      P33707;
DT      01-FEB-1994 (Rel. 28, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Erythropoietin precursor (Fragment).
GN      EPO.
OS      Canis familiaris (Dog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxID=9615;
RN      [1]
RP      SEQUENCE FROM N.A.,
RX      MEDLINE=93372347; PubMed=8364201;
RA      Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA      Czelusniak J., Goodman M., Bunn H.F.;
RT      "Erythropoietin structure-function relationships: high degree of
RT      sequence homology among mammals.";
RL      Blood 82:1507-1516(1993).
CC      -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC      PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC      -1- SUBCELLULAR LOCATION: Secreted.

```

```

CC      -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC      AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC      -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: L13027; AAA30842.1; -.
CC      DR      HSSP: P01588; 1CN4.
CC      DR      InterPro: IPR001323; EPO_TPO.
CC      DR      Pfam: PF00758; EPO_TPO; 1.
CC      DR      Prosite: PS00817; EPO_TPO; 1.
CC      KW      Erythrocyte maturation; Glycoprotein; Hormone; Signal.
CC      FT      NON_TER 1
CC      FT      SIGNAL <1 22
CC      FT      CHAIN 23 190
CC      FT      DISULFID 29 >175
CC      FT      DISULFID 51 55
CC      FT      CARBOHYD 46 46
CC      FT      CARBOHYD 60 60
CC      FT      CARBOHYD 105 105
CC      FT      CARBOHYD 175 175
CC      SQ      SEQUENCE 175 AA; 19193 MW; B504F8D8667B6F4 CRC64;

Query Match
Best Local Similarity 81.0%; Score 638; DB 1; Length 175;
Matches 124; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

OY      1 APPRLICDSRVLEERYLLLEKAEKNITTCGAEHCSLNENITVPDTKVFYAMKMEYGOQA 60
        |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB      23 APPRLICDSRVLEERYLLLEKAEKNITTCGAEHCSLNENITVPDTKVFYAMKMEYGOQA 82
OY      61 VEWOGALLSEAVLNGQALLVNSQWPEPLQHDVKAVSGLSLTLTLRALGAQKEAIS 120
        :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB      83 LEWOGALLSEAVLNGQALLVNSQWPEPLQHDVKAVSGLSLTLTLRALGAQKEAIS 142
OY      121 PPDAASAAPLRTTADTFKRLFRYSNPLRGKLTLYGECRCRD 153
        |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB      143 LPEASAPLRTTADTFKRLFRYSNPLRGKLTLYGECRCRD 175

RESULT 11
TPO_CANFA
ID      TPO_CANFA          STANDARD;          PRT;          352 AA.
AC      P42705;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
DE      (C-MPL ligand) (ML) (Megakaryocyte growth and development factor)
GN      (MGDF).
OS      Canis familiaris (Dog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxID=9615;
RN      [1]
RP      SEQUENCE FROM N.A., AND SEQUENCE OF 24-44.
RX      MEDLINE=94291201; PubMed=8020099;
RA      Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,
RA      Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
RA      Hsu R.-Y., Parker V.P., Sugas S., Skrine J.D., Werewether L.A.,
RA      Clogson C., Hsu E., Hukom M.M., Hornkohl A., Choi E., Pangelinan M.,
RA      Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C.,
RA      Shuter J., Chute N., Basu R., Selander L., Trollinger D., Siew L.,
RA      Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
RA      Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,

```

RA Pacifici R., Ponting I., Saris C., Wen D., Yung Y.P., Lin H.,
 RA Bosselman R.A.;
 RT "Identification and cloning of a megakaryocyte growth and development
 RT factor that is a ligand for the cytokine receptor Mpl.";
 RL Cell 77:1117-1117(1994).
 CC -1- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
 CC ACTIVATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
 CC CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
 CC MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DOMAIN: TWO-DOMAIN STRUCTURE. WITH AN ERYTHROPOIETIN-LIKE N-
 CC TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
 CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 DR InterPro: IPR001323; EPO_TPO.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PRINTS: PR01485; THROMBOPTN.
 DR PROSITE: PS00817; EPO_TPO; 1.
 KW Cytokine; Glycoprotein; Hormone; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 352
 FT DISULFID 28 172
 FT POTENTIAL.
 FT CARBOHYD 50 106
 FT N-LINKED (GLCNAC: . . .) (POTENTIAL):
 FT CARBOHYD 185 185
 FT N-LINKED (GLCNAC: . . .) (POTENTIAL):
 FT CARBOHYD 197 197
 FT N-LINKED (GLCNAC: . . .) (POTENTIAL):
 FT CARBOHYD 206 206
 FT N-LINKED (GLCNAC: . . .) (POTENTIAL):
 FT CARBOHYD 236 236
 FT N-LINKED (GLCNAC: . . .) (POTENTIAL):
 FT CARBOHYD 255 255
 FT N-LINKED (GLCNAC: . . .) (POTENTIAL):
 FT CARBOHYD 332 332
 FT N-LINKED (GLCNAC: . . .) (POTENTIAL):
 FT CARBOHYD 347 347
 FT N-LINKED (GLCNAC: . . .) (POTENTIAL):
 SO SEQUENCE 352 AA; 37641 MW; 024F3B41B061EDB8 CRC64;
 Query Match 12.9%; Score 109; DB: 1; Length 352;
 Best Local Similarity 24.0%; Pred. No. 0.0025;
 Matches 40; Conservative 23; Mismatches 62; Indels 42; Gaps 5;
 OY 1 APPRLCDSHVLEKLEAKENITTCGAEHCSLMENTVPTKXFAKRMMEGCOO 60
 DB 24 APP-ACDPRLLNKMRLDSHVLSRISQCPDIPSTPLPFAVDSELCGEMTKORQKA 81
 OY 61 VEWQGLALISEAVL--RQOALLVNSSQWPEPLQLHVDKAVGSLRSITTLRLALGQKKA 118
 DB 82 QDVGAVALLDGLVLAARQL-----GPSCLSLILGQLSGOYRL 120
 OY 119 I-----SPDPAASAPLRITTTADFERKLFERYVSNFLRGKIK 154
 DB 121 LIGALGGLGTQLPFG-----KTTTKDPNPAIFLPSFOQLRGKVR 161
 RESULT 12
 TPO_HUMAN STANDARD; PRT; 353 AA.
 AC P40325; Q13020; Q15790; Q15791; Q15792;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
 DE (Myelopoietin/leukemia virus oncogene ligand) (C-mpl ligand)
 DE (Mpl) (Megakaryocyte growth and development factor) (MGDF).
 GN THPO.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_Taxid:9606;
 RA de Saude Federal 11ver;
 RA MEDLINE-94261202; PubMed-8202154;
 RA Spencer S.A.F., J., Hass P.E., Spencer S.D., Malloy B.E., Gurney A.L.,
 RA Oles K.J., Hillier B., Solberg L.A., Jr., Goeddel D.V., Eaton D.L.;
 RT "Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-mpl
 RT ligand.";

RL Nature 369:533-538(1994).
 RL 1
 RL SEQUENCE FROM N.A. (ISOFORM 1).
 RP TISSUE-Fetal liver;
 RX MEDLINE-94291201; PubMed-8020099;
 RA Bartley T.P., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,
 RA Chang M.-S., Samal B.B., Nichols J., Sklar A.D., Chow E., Metwether L.A.,
 RA Hsu R.-Y., Parker V.P., Suggs S., Hornkohl A., Peniguelin M.,
 RA Clugson C., Hsu E., Hokom M.M., Hornkohl A., Peniguelin M.,
 RA Sun Y., Mar Y., McNinch J., Simonet L., Jacobsen F., Stau L.,
 RA Shutter J., Chute H., Basu R., Selander L., Trollinger D., Stau L.,
 RA Padilla D., Trail G., Elliott G., Izumi R., Covey T., Covey J.,
 RA Garcia A., Xu W., del Castillo J., Birn J., Cole S., Hu M.C.-T.,
 RA Pacifici R., Ponting I., Saris C., Wen D., Yung Y.P., Lin H.,
 RA Bosselman R.A.;
 RT "Identification and cloning of a megakaryocyte growth and development
 RT factor that is a ligand for the cytokine receptor Mpl.";
 RL Cell 77:1117-1124(1994).
 CC -1- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
 CC ACTIVATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
 CC CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
 CC MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2/TPO-2 AND
 CC 3)
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 RP Fetus;
 RX MEDLINE-95108091; PubMed-7809166;
 RA Foster D.C., Brecher C.A., Grant F.J., Kramer J.M., Kuipjer J.L.,
 RA Holly R.D., Brecher C.A., Grant F.J., Kramer J.M., Kuipjer J.L.,
 RA McGraw V., Hart C., Hargreave J., Heipel M.D., Bell L.A.N., Ching A.F.,
 RA "Human thrombopoietin: gene structure, cDNA sequence, expression, and
 RA chromosomal localization".
 RA Proc. Natl. Acad. Sci. U.S.A. 91:13023-13027(1994).
 CC [4]
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 RP Fetus;
 RX MEDLINE-95010765; PubMed-7926023;
 RA Sohma Y., Akahori H., Seki N., Hori T.-A., Ogami K., Kawamura K.,
 RA Miyazaki H.;
 RT "Molecular cloning and chromosomal localization of the human
 RT thrombopoietin gene.";
 RL FEBS Lett. 353:57-61(1994).
 CC [5]
 CC SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RP Fetus;
 RX MEDLINE-95152076; PubMed-7849319;
 RA Gurney A.L., Kuang W.-C., Xie M.-H., Malloy B.E., Eaton D.L.,
 RA "Genomic organization, chromosomal localization, and conserved
 RT alternative splice forms of thrombopoietin.";
 RL Blood 85:981-988(1995).
 CC [6]
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 RP Tissue-Liver;
 RX MEDLINE-96015174; PubMed-8537317;
 RA Kato T., Ogami K., Shimada Y., Iwamatsu A., Sohma Y., Akahori H.,
 RA Hori T., Ogami K., Shimada Y., Iwamatsu A., Sohma Y., Akahori H.,
 RA Ozawa T., Inoue H., Kawamura K., Miyazaki H.,
 RA "Purification and characterization of thrombopoietin.";
 RL J. Biochem. 118:229-236(1995).
 CC [7]
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 RP Tissue-Placenta;
 RX MEDLINE-95122483; PubMed-7822271;
 RA Chang M., McNinch J., Basu R., Shutter J., Hsu R., Perkins C., Mar V.,
 RA Suggs S., Welcher A., Li L., Lu H., Bartley T., Hunt P., Martin F.,
 RA Samal B., Bogenberger J.;
 RT "Cloning and characterization of the human megakaryocyte growth and
 RT development factor (MGDF) gene.";
 RL J. Biol. Chem. 270:511-514(1995).
 CC [8]
 CC SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
 RP Tissue-Fetal liver;
 RX MEDLINE-94261202; PubMed-8202154;
 RA de Saude Federal 11ver;
 RA MEDLINE-94261202; PubMed-8202154;
 RA Spencer S.A.F., J., Hass P.E., Spencer S.D., Malloy B.E., Gurney A.L.,
 RA Oles K.J., Hillier B., Solberg L.A., Jr., Goeddel D.V., Eaton D.L.;
 RT "Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-mpl
 RT ligand.";

```

RESULT 13
POLG_HCVJ8      STANDARD;          PRT;    3033 AA.
AC      P2661;
DT      01-AUG-1992 (Rel. 23, Created)
DT      01-AUG-1992 (Rel. 23, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Genome polypeptide [contains: Capsid protein C (core protein) (P22);
DE      Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
DE      (GP68) (NS1); Protein P7; Nonstructural protein NS2 (P21)
DE      (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirus)
DE      (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
DE      NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
DE      NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
OS      Hepatitis C virus (isolate HC-J8) (HCV).
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC      Hepacivirus.
OX      NCBI_Taxid=11115;
XX      1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=92230232; PubMed=1314459;
RA      Okamoto H., Kuzai K., Okada S.-I., Yamamoto K., Iizuka H., Tanaka T.,
RA      Fukuda S., Tsuda F., Mishiro S.;
RT      "Full-length sequence of a hepatitis C virus genome having poor
RT      homology to reported isolates: comparative study of four distinct
RT      genotypes."
RL      Virology 188:331-341(1992).
RC      -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
CC      HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
CC      NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC      -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
CC      precursor polypeptide, commonly with Asp or Glu in the P6
CC      position, Cys or Thr in P1 and Ser or Ala in P1'.
CC      -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC      (RNA)(N).
CC      -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC      LIPIDPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS.
CC      PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC      PROTEIN C AND RNA.
CC      -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; D10988; BAA01761.1; -.
DR      PIR; A40250; GNMVJ8.
DR      HSSP; P27958; 1HEI.
DR      MEROPS; S29_001; -.
DR      MEROPS; U39_001; -.
DR      InterPro; IPR001410; DEAD.
DR      InterPro; IPR002531; HCV_NS1.
DR      InterPro; IPR002518; HCV_NS2.
DR      InterPro; IPR004109; HCV_NS3.
DR      InterPro; IPR000745; HCV_NS4A.
DR      InterPro; IPR001490; HCV_NS4B.
DR      InterPro; IPR002868; HCV_NS5A.
DR      InterPro; IPR002166; HCV_RdRP.
DR      InterPro; IPR002522; HCV_capsid.
DR      InterPro; IPR002521; HCV_core.
DR      InterPro; IPR002519; HCV_env.
DR      Pfam; PF00998; HCV_RdRP; 1.
DR      Pfam; PF01001; HCV_NS4B; 1.
DR      Pfam; PF01006; HCV_NS4A; 1.
DR      Pfam; PF01506; HCV_NS5A; 1.

```

ID	Accession	Standard	Prot.	897 AA
Dr	PF01538	HCV NS2	1	
Dr	PF01539	HCV env	1	
Dr	PF01543	HCV core	1	
Dr	PF01543	HCV capsid	1	
Dr	PF01543	HCV NS1	1	
Dr	PF01543	HCV NS3	1	
Dr	PF01543	HCV NS4	1	
Dr	PF01543	HCV NS5	1	
Dr	PF01543	HCV NS6	1	
Dr	PF01543	HCV NS7	1	
Dr	PF01543	HCV NS8	1	
Dr	PF01543	HCV NS9	1	
Dr	PF01543	HCV NS10	1	
Dr	PF01543	HCV NS11	1	
Dr	PF01543	HCV NS12	1	
Dr	PF01543	HCV NS13	1	
Dr	PF01543	HCV NS14	1	
Dr	PF01543	HCV NS15	1	
Dr	PF01543	HCV NS16	1	
Dr	PF01543	HCV NS17	1	
Dr	PF01543	HCV NS18	1	
Dr	PF01543	HCV NS19	1	
Dr	PF01543	HCV NS20	1	
Dr	PF01543	HCV NS21	1	
Dr	PF01543	HCV NS22	1	
Dr	PF01543	HCV NS23	1	
Dr	PF01543	HCV NS24	1	
Dr	PF01543	HCV NS25	1	
Dr	PF01543	HCV NS26	1	
Dr	PF01543	HCV NS27	1	
Dr	PF01543	HCV NS28	1	
Dr	PF01543	HCV NS29	1	
Dr	PF01543	HCV NS30	1	
Dr	PF01543	HCV NS31	1	
Dr	PF01543	HCV NS32	1	
Dr	PF01543	HCV NS33	1	
Dr	PF01543	HCV NS34	1	
Dr	PF01543	HCV NS35	1	
Dr	PF01543	HCV NS36	1	
Dr	PF01543	HCV NS37	1	
Dr	PF01543	HCV NS38	1	
Dr	PF01543	HCV NS39	1	
Dr	PF01543	HCV NS40	1	
Dr	PF01543	HCV NS41	1	
Dr	PF01543	HCV NS42	1	
Dr	PF01543	HCV NS43	1	
Dr	PF01543	HCV NS44	1	
Dr	PF01543	HCV NS45	1	
Dr	PF01543	HCV NS46	1	
Dr	PF01543	HCV NS47	1	
Dr	PF01543	HCV NS48	1	
Dr	PF01543	HCV NS49	1	
Dr	PF01543	HCV NS50	1	
Dr	PF01543	HCV NS51	1	
Dr	PF01543	HCV NS52	1	
Dr	PF01543	HCV NS53	1	
Dr	PF01543	HCV NS54	1	
Dr	PF01543	HCV NS55	1	
Dr	PF01543	HCV NS56	1	
Dr	PF01543	HCV NS57	1	
Dr	PF01543	HCV NS58	1	
Dr	PF01543	HCV NS59	1	
Dr	PF01543	HCV NS60	1	
Dr	PF01543	HCV NS61	1	
Dr	PF01543	HCV NS62	1	
Dr	PF01543	HCV NS63	1	
Dr	PF01543	HCV NS64	1	
Dr	PF01543	HCV NS65	1	
Dr	PF01543	HCV NS66	1	
Dr	PF01543	HCV NS67	1	
Dr	PF01543	HCV NS68	1	
Dr	PF01543	HCV NS69	1	
Dr	PF01543	HCV NS70	1	
Dr	PF01543	HCV NS71	1	
Dr	PF01543	HCV NS72	1	
Dr	PF01543	HCV NS73	1	
Dr	PF01543	HCV NS74	1	
Dr	PF01543	HCV NS75	1	
Dr	PF01543	HCV NS76	1	
Dr	PF01543	HCV NS77	1	
Dr	PF01543	HCV NS78	1	
Dr	PF01543	HCV NS79	1	
Dr	PF01543	HCV NS80	1	
Dr	PF01543	HCV NS81	1	
Dr	PF01543	HCV NS82	1	
Dr	PF01543	HCV NS83	1	
Dr	PF01543	HCV NS84	1	
Dr	PF01543	HCV NS85	1	
Dr	PF01543	HCV NS86	1	
Dr	PF01543	HCV NS87	1	
Dr	PF01543	HCV NS88	1	
Dr	PF01543	HCV NS89	1	
Dr	PF01543	HCV NS90	1	
Dr	PF01543	HCV NS91	1	
Dr	PF01543	HCV NS92	1	
Dr	PF01543	HCV NS93	1	
Dr	PF01543	HCV		

```

DE Epidemic growth factor receptor substrate 15 (Protein Eps15) (AF-IP
GN Eps15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxId=10090;
RX MEDLINE=93361014; PubMed=7689153;
RP TISSUE=Fibroblast;
RC Fazio I F., Millicheolo L., Matoskova B., Wong W.T., di Fiore P.P.;
RT "Eps15, a novel tyrosine kinase substrate, exhibits transforming
RT activity."
RL MOL. Cell. Biol. 13:5814-5828(1993).
CC -1- FUNCTION: INVOLVED IN CELL GROWTH REGULATION. MAY BE INVOLVED IN
CC THE REGULATION OF MITOGENIC SIGNALS AND CONTROL OF CELL
CC PROLIFERATION.
CC -1- SUBUNIT: Interacts with Stc1n 2.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- PTM: PHOSPHORYLATED ON TYROSINE BY EGFR.
CC -1- SIMILARITY: CONTAINS 3 EH DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR HMBL: L21768; AAA02912.1; -.
DR HSSP: P02649; IBZ4.
DR MGD: MG1:104583; Ephs15.
DR InterPro: IPRO02048; EF-hand.
DR InterPro: IPRO00261; Eps15_repeat.
DR InterPro: IPRO03903; UIM.
DR Pfam: PF00036; efhand; 4.
DR SMART: SM00054; Eph; 3.
DR SMART: SM00027; EH; 3.
DR ProSITE: PS00018; EF_HAND; 2.
DR ProSITE: PS50031; EH; 3.
KW Phosphorylation; Calcium-binding; Repeat.
FT DOMAIN 15 104 EH 1.
FT CA_BIND 128 216 EH 2.
FT CA_BIND 173 184 EF-HAND 1 (POTENTIAL).
FT DOMAIN 224 314 EH 3.
FT CA_BIND 236 247 EF-HAND 2 (POTENTIAL).
FT DOMAIN 599 829 15 X 3 AA REPEATS OF D-E-F.
FT REPEAT 599 601 1
FT REPEAT 623 625 2.
FT REPEAT 629 631 2.
FT REPEAT 632 634 2.
FT REPEAT 636 638 2.
FT REPEAT 640 642 4.
FT REPEAT 645 647 5.
FT REPEAT 651 653 6.
FT REPEAT 653 655 7.
FT REPEAT 657 659 8.
FT REPEAT 693 695 9.
FT REPEAT 711 713 10.
FT REPEAT 713 715 11.
FT REPEAT 806 808 12.
FT REPEAT 827 829 13.
FT MOD_RES 132 132
SQ SEQUENCE 897 AA; 98470 MW; 0BA0CDD23FB73C2 CRC64;
      (PHOSPHORYLATION (BY TYR-KINASES))
Query Match          9.3%; Score 78.5; DB 1; Length 897;
Best Local Similarity 26.5%; Pred. NO. 8.2;
Matches 26; Conservative 20; Mismatches 35; Indels 17; Gaps 3;
12 LERYLLAEKKEKNITTTGCAEHHC---SLNENIVYDPDKRVNFYAWKKREVGQAQAVEVWGL 67
| | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Db 408 LEOQLQEVRRK-----CAEQAQLISLKAETISQESQISSYEELKLARELSRLQOET 461
 QY 68 ALISEAVLRGALLVNSSQPWEPIQLHYDKAVSGLRSL 105
 Db 462 AOLSESVESGKRAQL-----EPILOQHLQESQOEISSM 492

RESULT 15
 ATCS_SYNP7
 ID ATCS_SYNP7 STANDARD; PRT; 747 AA.
 AC P37279;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cation-transporting ATPase pacs (EC 3.6.3.-).
 GN PACS.
 OS Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 OX NCBI_TaxId=1140;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93380581; PubMed=8370468;
 RA Kanamaru K., Kashiwagi S., Mizuno T.;
 RT "The cyanobacterium, Synechococcus sp. PCC7942, possesses two
 RT distinct genes encoding cation-transporting P-type ATPases.";
 RL FEBS Lett. 330:99-104(1993).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE OSMOTIC ADAPTATION.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES). SUBFAMILY IB.
 CC -1- SIMILARITY: CONTAINS 1 HMA DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D16437; BAA03907.1; -.
 DR PIR: S36741; S36741.
 DR HSSP: 004656; 1A00.
 DR InterPro: IPR001757; ATPase_E1-E2.
 DR InterPro: IPR001934; HeavyMg_transp.
 DR InterPro: IPR001454; Hlgase/Hydrlase.
 DR Pfam: PF00122; E1-E2_ATPase; 1.
 DR Pfam: PF00403; HMA; 1.
 DR Pfam: PF00702; Hydrolase; 1.
 DR PROSITE: PS00154; ATPASE_E1-E2; 1.
 DR PROSITE: PS01047; HMA_1; 1.
 DR PROSITE: PS50846; HMA_2; 1.
 DR HydroLase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;
 KW Metal-binding.
 FT DOMAIN 1 101 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 102 122 POTENTIAL.
 FT DOMAIN 123 132 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 133 151 POTENTIAL.
 FT DOMAIN 152 158 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 159 179 POTENTIAL.
 FT DOMAIN 180 199 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 200 220 POTENTIAL.
 FT DOMAIN 221 348 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 349 371 POTENTIAL.
 FT DOMAIN 372 378 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 379 396 POTENTIAL.
 FT DOMAIN 397 688 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 689 708 POTENTIAL.
 FT DOMAIN 709 720 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 721 739 POTENTIAL.
 FT DOMAIN 740 747 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 740 747 HMA.

FT MOD_RES 434 434 PHOSPHORYLATION (BY SIMILARITY).
 FT METAL 14 14 POTENTIAL.
 FT METAL 17 17 POTENTIAL.
 FT METAL 634 634 MAGNESIUM (BY SIMILARITY).
 FT METAL 638 638 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 747 AA; 79732 MW; 0AF148B92EB85CEE CRC64;

Query Match 9.2%; Score 78; DB 1; Length 747;
 Best Local Similarity 25.4%; Pred. No. 7.3;
 Matches 33; Conservative 22; Mismatches 45; Indels 30; Gaps 5;

QY 17 LEAKAEANTTGCARHCSTENITVDPFKVNFYANKRMVGGOAV---EVMGGLALLSRA 73
 Db 233 LQAKTARVLRQG-----QELTLPITVEQYEDWVRVPPGKVPYDGEVIGRSTVDES 284
 QY 74 VLRGALLVNSSQPWEPIQLHYDKAVSGLR-----SLTTLRLALG-----AQKEAISPPD 123
 Db 285 MVTGESL-----PVQKQYGVDEYIGATLTKTGSITIRATRVGRHETFLAQIVQLVQQA 335
 QY 124 AASAPLRTI 133
 Db 336 QASKAPLQRL 345

Search completed: January 2, 2003, 15:11:51
 Job time : 10.4743 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:09:40 : Search time 26.4199 Seconds
(without alignments) updates/sec
1286.823 Million cell

Title: US-09-853-731-1

Perfect score: 846
Sequence: 1 APPRLICDSRVLEKAEYTLLEAK.....SNFLGSKLKYTGACRTGD 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	680.5	80.4	195	6 O9GKA3	O9gka3 oryctolagus
2	680.5	80.4	195	6 O9GKA2	O9gka2 oryctolagus
3	678	80.1	194	6 O9MYM8	O9mym8 sus scrofa
4	188	22.2	50	11 O9OV40	O9ov40 rattus sp.
5	88	10.4	323	16 O8ZDC8	O8zdc8 yersinia pe
6	87.5	10.3	346	16 O8ZKZ4	O8zkz4 salmonella
7	87.5	10.3	346	16 O8Z2M5	O8z2m5 salmonella
8	85	10.0	3722	2 P94873	P94873 lysobacter
9	83	9.8	296	16 O8ZAY4	O8zay4 yersinia pe
10	83	9.8	333	16 O9HZM7	O9hzm7 pseudomonas
11	82.5	9.8	303	12 O9DH06	O9dh06 hepatitis c
12	82	9.7	815	10 O9EK91	O9ek91 arabidopsis
13	79.5	9.4	1829	16 O86808	O86808 streptomyc
14	79.5	9.4	3033	12 O91ZAI	O91zai hepatitis c
15	79	9.3	480	10 O43380	O43380 avena sativ
16	79	9.3	2364	5 Q22896	Q22896 caenorhabdt

17	78.5	9.3	813	16 O8Z9B9	O8z9b9 salmonella
18	78.5	9.3	3019	12 O92529	O92529 hepatitis c
19	78	9.2	348	16 O86751	O86751 streptomyc
20	78	9.2	378	17 O8ZK3	O8zk3 pyrobaculum
21	78	9.2	545	16 O8UB09	O8ub09 agrobacteri
22	78	9.2	458	2 O93748	O93748 buchnera ap
23	78	9.2	548	2 O93N35	O93n35 buchnera ap
24	77.5	9.2	242	16 O8Y76	O8y76 anabaena sp
25	77.5	9.2	310	16 O91033	O91033 streptomyc
26	77.5	9.2	451	16 O8XSE8	O8xse8 raietonia s
27	77.5	9.2	1003	10 O9FIR4	O9fir4 arabidopsis
28	77.5	9.2	2556	11 O91XW2	O91xw2 mus musculu
29	77	9.1	379	2 O9L9E5	O9l9e5 streptomyc
30	77	9.1	383	16 O9ZOV7	O9zov7 rhizobium m
31	77	9.1	447	2 P72270	P72270 rhodococcus
32	77	9.1	451	16 P74054	P74054 synecocyst
33	76.5	9.0	154	16 O9PGAT	O9pgat xylella tas
34	76.5	9.0	425	16 O8YF20	O8yf20 bruceella me
35	76.5	9.0	619	3 O96VB7	O96vb7 botrytis cl
36	76.5	9.0	1315	3 O8X215	O8x215 botrytis cl
37	76.5	9.0	1315	3 O8X1E7	O8x1e7 botrytis cl
38	75.5	8.9	637	16 O5413	O5413 synecocyst
39	75.5	8.9	871	5 O9NB04	O9nb04 drosophila
40	75	8.9	308	10 O94141	O94141 oryza sativ
41	75	8.9	397	8 O9TNG5	O9tng5 coriaria te
42	75	8.9	548	2 O93N34	O93n34 buchnera ap
43	75	8.9	554	2 O9RPH5	O9rph5 mycobacteri
44	74.5	8.8	361	6 O9T5W6	O9tsw6 sus scrofa
45	74.5	8.8	389	2 O9AN18	O9an18 bradyrhizob

ALIGNMENTS

RESULT 1	
O9GKA3	PRELIMINARY; PRT; 195 AA.
AC O9GKA3:	
DT 01-MAR-2001 (TREMBLrel. 16, Created)	
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)	
DE Erythropoietin.	
OS Oryctolagus cuniculus (Rabbit).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.	
OX NCBI_TaxId=9986;	
RN [1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE=21290682; PubMed=11396976;	
RA Villalta A., Wu D., Margalith M., Hobart P.;	
RT "Rabbit EPO Gene and CDNA: Expression of Rabbit EPO after	
RT Intramuscular Injection of pDNA.";	
RL Biochem. Biophys. Res. Commun. 284:823-827(2001).	
DR EMBL; AF290943; AAG36961.1; -.	
DR HSSP; P01588; 1CN4.	
DR InterPro; IPR001323; EPO.TPO.	
DR InterPro; IPR003013; Erythropo.	
DR Pfam; PF00758; EPO.TPO.1.	
DR PRINTS; PR00272; ERYTHROPTN.	
DR PROSITE; PS00817; EPO.TPO.1.	
SQ SEQUENCE 195 AA; 21053 MW; 0999DA7D852713F3 CRC64;	
Query Match	80.4%; Score 680.5; DB 6; Length 195;
Best Local Similarity	81.3%; Pred. No. 1,4e-60;
Matches 195; Conservative 12; Mismatches 18; Indels 1; Gaps 1;	
OY 1 APPRLICDSRVLEKAEYTLLEAKAEYITGCAEHCSLNTNITVPDKNVPMYAKRMVEVGOA 60	
DB 29 APPRLICDSRVLEKAEYTLLEAKAEYITGCAEHCSLNTNITVPDKNVPMYAKRMVEVGOA 88	
OY 61 VEWOGALALSEAVLRGQALLVNSOSPWEPTOLVHDKAVSGSLRTTLRLALGAKGKAIS 120	
DB 89 VEWOGALALSEAVLRGQALLVNSOSPWEPTOLVHDKAVSGSLRTTLRLALGAKGKAIS 148	

OY 121 PPDA--SAAPLRTTADTFKRLFRVYSNPLRGKILYTGACRGTG 165
 DB 149 PEPASSAAPLRTVADTLCIKLFRYSNPLRGKILYTGACRGTG 194

RESULT 2

O9GKA2 PRELIMINARY; PRT; 195 AA.
 AC O9GKA2: 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 OS Erythropoietin
 OC Oryctolagus cuniculus (Rabbit)
 CC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21290682; PubMed=11396976;
 RA Vilalta A., Wu D., Margalit M., Hobart P.,
 RT "Rabbit EPO Gene and cDNA: Expression of Rabbit EPO after
 RT Intramuscular Injection of pDNA"
 RL Biochem. Biophys. Res. Commun. 284:823-827(2001).
 DR EMBL; AF290944; AAC36962.1; ..
 DR HSSP; P01588; 1CN4.
 DR InterPro; IPR001323; EPO_TPO.
 DR Pfam; PF00758; EPO_TPO; 1.
 DR PRINTS; PR00272; ERYTHROPTN.
 DR PROSITE; PS00817; EPO_TPO; 1.
 SQ SEQUENCE 195 AA; 21025 MW; 1F1DC7F403A303BC CRC64;

Query Match 80.4%; Score 680.5; DB 6; Length 195;
 Best Local Similarity 81.3%; Pred. No. 1.4e-60;
 Matches 133; Conservative 12; Mismatches 18; Indels 1; Gaps 1;

OY 1 APRLLDSDRYLERYLLEAKAEENITTCGAHCSLNENITVPDTKYNFYAMKREYGOOA 60
 DB 29 APAPLLDSDRYLERYLLEAKAEENITTCGAHCSLNENITVPDTKYNFYAMKREYGOOA 88
 OY 61 VEWOGIALISEAVLRGOALLVNSQWPEPLQIHDVKAVSGLSLTTLRALGAOKEAIS 120
 DB 89 VEWOGIALISEAVLRGOALLVNSQWPEPLQIHDVKAVSGLSLTTLRALGAOKEAIS 148
 OY 121 PPDA--SAAPLRTTADTFKRLFRVYSNPLRGKILYTGACRGTG 165
 DB 149 PEPASSAAPLRTVADTLCIKLFRYSNPLRGKILYTGACRGTG 194

RESULT 3

O9MYM8 PRELIMINARY; PRT; 194 AA.
 AC O9MYM8: 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 OS Erythropoietin precursor.
 CC Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OX Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 RN NCBI_TaxID=9823;
 RP SEQUENCE FROM N.A.
 RC STRAIN=NORWEGIAN LANDRACE; TISSUE=KIDNEY;
 RA David B. Harbitt, J.,
 RT "The porcine erythropoietin gene: cDNA and genomic sequences and
 RT expressed (SPP-1993)."
 RL EMBL; AJ249745; CAB96416.1; ..
 DR EMBL; AJ249745; CAB96417.1; ..

DR HSSP; P01588; 1CN4.
 DR InterPro; IPR001323; EPO_TPO.
 DR InterPro; IPR001323; Erythropo. 1
 DR Pfam; PF00758; EPO_TPO; 1
 DR PRINTS; PR00272; ERYTHROPTN.
 DR PROSITE; PS00817; EPO_TPO; 1.
 KW SIGNAL.
 FT CHAIN 1 26 POTENTIAL.
 FT CHAIN 27 194 ERYTHROPTENIN.
 SQ SEQUENCE 194 AA; 21303 MW; 77861A6F62DEA1C CRC64;

Query Match 80.1%; Score 678; DB 6; Length 194;
 Best Local Similarity 82.0%; Pred. No. 2.5e-60;
 Matches 137; Conservative 7; Mismatches 21; Indels 2; Gaps 1,
 OY 1 APRLLDSDRYLERYLLEAKAEENITTCGAHCSLNENITVPDTKYNFYAMKREYGOOA 60
 DB 27 APAPLLDSDRYLERYLLEAKAEENITTCGAHCSLNENITVPDTKYNFYAMKREYGOOA 86
 OY 61 VEWOGIALISEAVLRGOALLVNSQWPEPLQIHDVKAVSGLSLTTLRALGAOKEAIS 120
 DB 87 VEWOGIALISEAVLRGOALLVNSQWPEPLQIHDVKAVSGLSLTTLRALGAOKEAIP 146
 OY 121 PPDA--SAAPLRTTADTFKRLFRVYSNPLRGKILYTGACRGTG 165
 DB 147 LPDAPSSAPLRTVADTLCIKLFRYSNPLRGKILYTGACRGTG 193

RESULT 4

O9GV40 PRELIMINARY; PRT; 50 AA.
 AC O9GV40: 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 OS Erythropoietin (Fragment).
 CC Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94115047; PubMed=7764337;
 RA Okano M., Suga H., Masuda S., Nagao M., Narita H., Ikura K.,
 RA Sasaki K.,
 RL Biosci. Biotechnol. Biochem. 57:1882-1885(1993).
 DR HSSP; P01588; 1EER.
 DR InterPro; IPR001323; EPO_TPO.
 DR Pfam; PF00758; EPO_TPO; 1.
 DR PRINTS; PR00272; ERYTHROPTN.
 SQ SEQUENCE 50 AA; 5587 MW; 70B44A8BFE016034 CRC64;

Query Match 22.2%; Score 188; DB 11; Length 50;
 Best Local Similarity 76.0%; Pred. No. 9.9e-12;
 Matches 35; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 OY 1 APRLLDSDRYLERYLLEAKAEENITTCGAHCSLNENITVPDTKYNFYA 50
 DB 1 APRLLDSDRYLERYLLEAKAEENITTCGAHCSLNENITVPDTKYNFYA 50

RESULT 5

O8ZDC8 PRELIMINARY; PRT; 323 AA.
 AC O8ZDC8: 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 OS Ribonucleoside-diphosphate reductase 2 beta chain (EC 1.17.4.1).
 GN NNDP OR YPO2648.
 OS Versinia pectis.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;


```

OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
MEDLINE=21470413; PubMed=11586360;
RA Parikh H.J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Fellwell T., Hamlin N., Holroyd S., Jagsels K., Karlyshev A.V.,
RA Leathers S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL, AF414153; CAC92889.1; -.
DR InterPro: IPR000358; RibonucL_redctse.
DR Pfam: PF00268; ribonuc_red_sm; 1.
DR PROSITE: PS00368; RIBORED_SMALL; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 323 AA; 36823 MW; 87C21F7BB9B7FD2 CRC64;

Query Match 10.4%; Score 88; DB 16; Length 323;
Best Local Similarity 25.2%; Pred. No. 1.3;
Matches 34; Conservative 20; Mismatches 59; Indels 22; Gaps 5;

OY 38 NITVPDTVNFYAKRMVEYGOAVVWOGALISEAVLRGQALVNSSQWPEPLQLHYD- 96
DB 2 NVKRTIRISALINMKRIE-DDKDELVWN--RLTSNFWLPEKVPISNDIPSWATLTPHEQO 58
OY 97 ---KAVSGLSRLTTLRALGAQ---KEAISPDAASAPLRTITDTFRKLFYVSNFLR 150
DB 59 LFIKFTGLTLDITQNTLGAPALIKDALTPEEAIFSNISMEVHAHSYSITSTL-- 116
OY 151 GKLLKTYGEACRTGD 165
DB 117 -----CLTSPD 121

RESULT 6
O82K24 PRELIMINARY; PRT; 346 AA.
AC O82K24;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Periplasmic sensor in multi-component regulatory system with Tors
DE (sensory kinase) and TorR (regulator), regulates tor operon.
GN TORR OR STM3825.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Lt2 / SGC1412 / ATCC 700720;
MEDLINE=21534948; PubMed=11677609;
RA McGettland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvoney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
DR EMBL, AE008878; AAL22684.1; -.
DR InterPro: IPR001761; PeriplabP/LacI.
DR Pfam: PF00532; Peripla_BP_like; 1.
KW Complete proteome.
SQ SEQUENCE 346 AA; 38440 MW; E37CAB58E49FD716 CRC64;

Query Match 10.3%; Score 87.5; DB 16; Length 346;
Best Local Similarity 26.7%; Pred. No. 1.5;

```

```

Matches 44; Conservative 22; Mismatches 48; Indels 51; Gaps 9;

OY 10 RYLERVYLEAKAEENITTG--CAEHCSLNE--NITVPDTKVFYAKRMVEYGOAVVWQ 65
DB 217 RNLQDEMLERHPDANVAVGSAIAEAAMGEGRNLTPTLTIVSFYL-----THQYVR 267
OY 66 GIALISEAVLRGQALVNSSQ--PWEPLQLHYDKAVSGLSLTTLRALGAQ--KEAISP 122
DB 268 GLK-----KHILMALSDQAMQ-----GELATQSIKIVLOGQVYPENISPP 309
OY 123 -----DASAPAPLRTITADTFRKLFYVSNFLRKGLKLYTGEA 160
DB 310 VLILTHNNADSARVRRSLSPGFRPVY-----LYQYTSSEA 344

RESULT 7
O82ZM5 PRELIMINARY; PRT; 346 AA.
AC O82ZM5;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Solute binding receptor protein.
GN STY3952.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
MEDLINE=21534947; PubMed=11677608;
RA Parikh H.J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Fellwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagsels K.,
RA Krogh A., Larsen T.S., Leathers S., Moule S., O'Gaora P., Barry C.,
RA Quail M., Rutherford K., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
DR EMBL, AL627280; CAD03169.1; -.
DR InterPro: IPR001761; PeriplabP/LacI.
DR Pfam: PF00532; Peripla_BP_like; 1.
KW Receptor; Complete proteome.
SQ SEQUENCE 346 AA; 38546 MW; F80FB168B8C3A8F CRC64;

Query Match 10.3%; Score 87.5; DB 16; Length 346;
Best Local Similarity 26.7%; Pred. No. 1.5;
Matches 44; Conservative 22; Mismatches 48; Indels 51; Gaps 9;

OY 10 RYLERVYLEAKAEENITTG--CAEHCSLNE--NITVPDTKVFYAKRMVEYGOAVVWQ 65
DB 217 RNLQDEMLERHPDANVAVGSAIAEAAMGEGRNLTPTLTIVSFYL-----THQYVR 267
OY 66 GIALISEAVLRGQALVNSSQ--PWEPLQLHYDKAVSGLSLTTLRALGAQ--KEAISP 122
DB 268 GLK-----KHILMALSDQAMQ-----GELATQSIKIVLOGQVYPENISPP 309
OY 123 -----DASAPAPLRTITADTFRKLFYVSNFLRKGLKLYTGEA 160
DB 310 VLILTHNNADSARVRRSLSPGFRPVY-----LYQYTSSEA 344

RESULT 8
P94873 PRELIMINARY; PRT; 3722 AA.
AC P94873;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

```

DE Alpha-aminoadipyl-cysteinyl-valine synthetase.
 GN PCBAB.
 OS Lysobacter lactamgenus.
 OC Bacteriia; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Lysobacter.
 NCBI_TaxID=39596;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KY90;
 RX MEDLINE=96275949; PubMed=8737573;
 RT Kimura H., Miyashita H., Sumino Y.;
 RT "Organization and expression in Pseudomonas putida of the gene cluster
 RT involved in cephalosporin biosynthesis from lysobacter lactamgenus
 RT KY90."
 RT APPL. Microbiol. Biotechnol. 45:490-501(1996).
 DR EMBL; D50308; BAA08846.1; -.
 DR HSSP; P14687; 1AMU.
 DR InterPro; IPR002106; AAcRNA_LigaseII.
 DR InterPro; IPR000873; AMP-bind.
 DR InterPro; IPR001242; Condensatn.
 DR InterPro; IPR000977; DNA_Ligase.
 DR InterPro; IPR003880; Pantine_attach.
 DR InterPro; IPR000379; Ser-ester_sile.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00501; AMP-binding; 3.
 DR Pfam; PF00668; Condensation; 3.
 DR Pfam; PF00550; PP-binding; 3.
 DR PROSITE; PS00179; AA-TRNA_LIGASE_II_1; UNKNOWN_1.
 DR PROSITE; PS00075; ACP_DOMAIN; 3.
 DR PROSITE; PS00455; AMP_BINDING; 1.
 DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_2.
 KW Phosphopantetheine.
 SQ SEQUENCE 3722 AA; 411607 MW; 3597B3483463809B CRC64;

Query Match Best Local Similarity 10.0%; Score 85; DB 2; Length 3722;
 Matches 34; Conservative 18; Mismatches 50; Indels 52; Gaps 5;
 QY 16 LLEAKENITTCGAHCISINENTVDPDKYFYAMKMEYGOQAVEY---WQGLALLS 71
 DB 2138 LLEGATQTLF-----PLPATYADFVWQROQLSDRDALPDYWG----- 2178
 QY 72 EAVLRGQALLVNSQPMPEIQLHVDKAVSGLSLTLTLRALGAQ-----KEAISPDAAS 126
 DB 2179 -----RSLAGQMPQLDLPD-----HARPAQFDYLGREIVFDVADAT 2214
 QY 127 AAPRTTADTFERKLFYVSNFLGKLLKYTGEA 160
 DB 2215 CDQLRVLAQTRTSFSEVLLAAVYLLTKAVSGS 2248

RESULT 9
 ID 08ZAY4 PRELIMINARY; PRT; 296 AA.
 AC 08ZAY4;
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-JUN-2002 (Tremblrel. 20, Last sequence update)
 DE Putative 2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60).
 GN Yersinia pestis.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 NCBI_TaxID=632;
 RP SEQUENCE FROM N.A.
 RC STRAIN=92/BIOVAR ORIENTALIS;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parshil J., Wren B.W., Thomson N.R., Tibbali R.W., Holden M.T.G.,
 RA Prentice H.B., Sepahnia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,

RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague";
 RL Nature 413:523-527(2001).
 DR EMBL; AJ414158; CAC93117.1; -.
 DR InterPro; IPR002204; 3hydroxisobut_dh.
 DR InterPro; IPR001744; 6PGD.
 DR Pfam; PF03446; NAD_binding_2; 1.
 DR PROSITE; PS00895; 3-HYDROXYISOBUT_DH; UNKNOWN_1.
 KW Oxidoreductase; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 296 AA; 32140 MW; C883AFCC6868429D CRC64;

Query Match Best Local Similarity 9.8%; Score 83; DB 16; Length 296;
 Matches 48; Conservative 34; Mismatches 62; Indels 54; Gaps 10;
 QY 6 ICDSRVLEKLEAK-EAENITTCGAHCISINENTVDPDKYFYAMKMEYGOQAVEY 63
 DB 32 LCDADITKVOITLTAELNAESITVDNA--ASIDLITLIPNSE--AVEQVLGSDGISG 85
 QY 64 WQGLALISEAVLRGQALLVNSQP-----WE-----PQLHVDKAVSG----- 101
 DB 86 W--VAQLSQAAV--VIDMSSSDPERSRLAILLAWEIDYLDAPVSGGVKKNQNTLSI 140
 QY 102 -----LNSLTTLRALGAQKEAISPDAASAP-----LRTTADTFERKLF 142
 DB 141 LIGGEDRYLKSCTYALAAWGEQILFVGPAAGSHAAALNNYSATGILITLALHVAQRF 200
 QY 143 RYVSNFLRGKLLKYTGEA 160
 DB 201 GIEPEVATEVLTNLTGNS 218

RESULT 10
 ID 09HZM7 PRELIMINARY; PRT; 339 AA.
 AC 09HZM7;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-JUN-2001 (Tremblrel. 16, Last sequence update)
 DE UDP-N-acetylglutamate decarboxylase.
 GN MURR OR PA2977
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 NCBI_TaxID=287;
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stoyer C.K., Phan X.Q., Ertan A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy S., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brady L.L., Coulter S., Rojter K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Hong G.K.S., Wang Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen";
 RL Nature 406:959-964(2000).
 DR EMBL; AE004723; AAG06385.1; -.
 DR HSSP; P08373; 2MBR.
 DR InterPro; IPR003170; MURR.
 DR InterPro; IPR001575; OXID_PAD_bind.
 DR InterPro; IPR000531; TOB_DoxC.
 DR Pfam; PF01565; PAD_binding_4; 1.
 DR Pfam; PF02873; MURR_C_1mg_4; 1.
 DR PROSITE; PS00430; TONE_DEPENDENT_REC_1; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 339 AA; 37627 MW; C0C8EF92938FE27 CRC64;
 Query Match 9.8%; Score 83; DB 16; Length 339;

Best Local Similarity 22.7%; Pred. No. 4.3;
Matches 41; Conservative 22; Mismatches 48; Indels 70; Gaps 8;

OY 31 EHCISLNE-NITVPTKVFYAKRMVEGQAVEWOGIALISEAVLRCQALLV----- 82
Db 7 EHCISLKPNTGIDVRAFLAHARDE-----ADVREALALARE--RCLPLVIGGCSNL 58
OY 83 -----NSSQPWEP-LQLHVCKAVSGLSRLTLL 109
Db 59 LITRDVEALVLRMASQGRIVSDAASVLAEEAGEAMDPEYQWSLREGALLENLSL- 117
OY 110 RALGAQKEAISPPDASAPLRTITA-----DTRKLFERYVSNPLRGKLTLYGEACRT 163
Db 118 -----PCTVGAAPMGNICAGVELKDVDFSLFAL--DQDGTILREFDQACRF 163
OY 164 G 164
Db 164 G 164

RESULT 11

OY 99DHD6 PRELIMINARY; PRT; 3033 AA.
AC Q9DHD6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Genome polypeptide (Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1))
DE Hepatitis C virus type 2b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31650;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21296595; PubMed=11402859;
RA Murakami K., Abe M., Kageyama T., Kamoshita N., Nomoto A.;
RT "Down-regulation of translation driven by Hepatitis C virus internal ribosomal entry site by the 3' untranslated region of RNA.";
RL Arch. Virol. 146:729-741(2001).
CC -1 SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR EMBL: AB030907; BAB08107.1.;
DR HSSP: p27958.1AIV.
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR002518; HCV_NS2.
DR InterPro: IPR004109; HCV_NS3.
DR InterPro: IPR000745; HCV_NS4.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_RdRp.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00998; HCV_RdRp; 1.
DR ProDom: PD186062; HCV_NS1; 1.
DR SMART: SM00487; DEXDC; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
KM Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KM Polypeptide; RNA-directed RNA polymerase; Transmembrane.
FT CHAIN 1 191 CORE PROTEIN.
FT CHAIN 192 383 E1 PROTEIN.

FT CHAIN 384 750 E2 PROTEIN.
FT CHAIN 751 813 P7 PROTEIN.
FT CHAIN 814 1030 NS2 PROTEIN.
FT CHAIN 1031 1661 NS3 PROTEINASE/HELICASE.
FT CHAIN 1662 1715 NS4A PROTEIN.
FT CHAIN 1716 1976 NS4B PROTEIN.
FT CHAIN 1977 2442 NS5A PROTEIN.
FT CHAIN 2443 3033 NS5B RNA-DEPENDENT RNA POLYMERASE.
SQ SEQUENCE 3033 AA; 329981 MW; 6B183FED090872B4 CRC64;

Query Match

Best Local Similarity 9.8%; Score 82.5; DB 12; Length 3033;
Matches 40; Conservative 21; Mismatches 60; Indels 27; Gaps 7;

OY 11 VIERYLEKAEENITTCGAE---HCSLNEITV-PDTKVFYAKRM-----EV 56
Db 1665 VLACGVLAAVAAYCLATGICISITIGRIHLNDQVAVAPDKELYEADMECEKALIBE 1724
OY 57 GQOAVEVWOG--LALLSEAVLRGQALLVNSSQPWEPLO----LHYDKAVSGLSRLTLLR 110
Db 1725 GQMAEMLSKLTGLLQOFTKQADIQPMQSSWPXIEQFMARHMMNFTSGIQYLAGLST 1784
OY 111 ALG---AQKEAISPPDASAPLRTIT 134
Db 1785 LPGNPAVASMAFS---AALTSPLEPTST 1809

RESULT 12

OY 99FK91 PRELIMINARY; PRT; 815 AA.
AC Q9FK91;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Seed maturation protein PM38 protein.
DE Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae: Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eutrosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned pl and TAC clones.";
RL DNA Res. 5:203-216(1998).
DR EMBL: AB012244; BAB09119.1.;
DR InterPro: IPR002086; Aldehyde_dehydr.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR001290; PARP.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF00644; PARP; 1.
DR SMART: SM00292; BRCT; 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE: PS50172; BRCT; 1.
SQ SEQUENCE 815 AA; 91534 MW; 4B1B602057D46B60 CRC64;

Query Match

Best Local Similarity 9.7%; Score 82; DB 10; Length 815;
Matches 42; Conservative 29; Mismatches 52; Indels 70; Gaps 12;

OY 21 EAENITTCGA-----EHCISLNEITVPTKVFYAKRMVEG-----QOAVEW 64
Db 328 EKDGLYNCAFSCIDLGKGRNEYC-IMQLVTPDSNLNMY-FKRGVGDPPNAEFLTEEM 385
OY 65 OGIALISEAVLRGQALLV-----NSSQPWEP-----PLQ-----LHYDKAVSGL 102
Db 386 ED-----EEAAIKEFARLEIAGNEFEPEWRBRKTIQKPKKFPIDMDGIEVRSGLGL 441

```

OY 103 RSL-----TTTLRALGAO-----KEAISPDDAASAPLRTTADTFRKL 141
DB 442 RQGIASAHCKLDSEFVANFIKVLGCEITNYALMELGLDPPD-----LPMGLMIDHLKRC 497
OY 142 FRYYSNPLNGKTK 154
DB 498 EEVLLEFVE-KVK 509

RESULT 13
O86808
ID O86808 PRELIMINARY; PRT: 1829 AA.
AC O86808:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative sensory histidine kinase.
GN SC05748 OR SC7C7.03.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_Taxid=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Harris D., Taylor K.;
RL Submitted (Jul-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (Jul-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP Submitted (Jul-1998) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA MEDLINE-97000351, PubMed-8843436;
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Khasn H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares R., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2) ";
RL Nature 417:141-147(2002).
CC -I STIMULATORY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
CC EMBL: AL031031; CA19849.1; -.
DR HSP; P52934; I23.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR003018; GAF.
DR InterPro: IPR003660; HAMF.
DR InterPro: IPR003661; His_kinA.
DR InterPro: IPR004359; His_kinA_sig.
DR InterPro: IPR001789; Response_reg.
DR Pfam: PF01590; GAF; 1.
DR Pfam: PF00672; HAMF; 11.
DR Pfam: PF02518; HATPase_C; 1.
DR Pfam: PF00072; response_reg; 1.
DR Pfam: PF00512; signal; 1.
DR PRINTS: PR00344; BCTRSENSOR.

```

```

DR ProDom: PD000039; Response_reg; 1.
DR SMART; SM00065; GAF; 1.
DR SMART; SM00304; HAMF; 12.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00388; HisKA; 1.
DR SMART; SM00448; REC; 1.
DR KINase; Phosphorylation; Sensory transduction; Transferase.
SQ SEQUENCE 1829 AA; 195751 MW; C9C869938C956A4 CRC64;

Query Match 9.4%; Score 79.5; DB 16; Length 1829;
Best Local Similarity 22.8%; Pred. No. 82;
Matches 37; Conservative 29; Mismatches 61; Indels 35; Gaps 6;

OY 18 EAKEENITTCGAHCSCINENITVPDKVNFYAMKRMVGGQAVEWQGLALL----- 70
DB 328 QVREISHTTAVA-NGDLSKRVTP-----ARGEVQLAETINQMTEFLIFADEV 377
OY 71 -----SEAVLRGOALLVNSQWPEPLQLHVDKAVSGLSRTITLRALGAOKEAISP 122
DB 378 TVANETGEGQLGGQAVPGAGIWKDL---TDSVNTVFNNITQYRDIAVTTAVASG 434
OY 123 DAASAPLRTTADTFRKLFVYS--NFLRGKALKYGEACR 162
DB 435 DIS-----QKVTVDVAGMELKNTVNTMDLSAFGAEVTR 471

RESULT 14
Q91ZAI
ID Q91ZAI PRELIMINARY; PRT: 3033 AA.
AC Q91ZAI:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Genome polypeptide [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)].
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_Taxid=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MD2B-1;
RA Itakura J., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,
RA Sato C.;
RT "Full length cDNA sequence of HCV genotype 2b, strain MD2b-1.";
RT Submitted (Feb-2000) to the EMBL/Genbank/DBJ databases.
CC -I STIMULATORY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR EMBL: AF388486; AF39945.1; -.
DR HSP; P27958; IHE1.
DR MEROPS; S29.001; -.
DR MEROPS; U39.001; -.
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002523; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_core.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR002518; HCV_NS2.
DR InterPro: IPR004109; HCV_NS3.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_RdRP.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01339; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02007; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00998; HCV_RdRP; 1.

```

DR PRODOM: PD186062; HCV NS1. 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; RNA-directed RNA polymerase; Transmembrane.
 SQ SEQUENCE 3033 AA; 330723 MW; FE04FEC7C385A13A CRC64;

Query Match 9.4%; Score 79.5; DB 12; Length 3033;
 Best Local Similarity 26.6%; Pred. No. 1.6e+02;
 Matches 41; Conservative 22; Mismatches 64; Indels 27; Gaps 7;

OY 5 LICDSRVLEKYLEAKEENITTCGAE--HCSLNENITV--PDTKVFNFYAMKRM----- 54
 DB 1659 IWTSSWVLAGGLAVAVAYCATGICISITGRHLNDQYVAPDKRILYEAFFDEMECASK 1718
 OY 55 ----EVGQAAVEWOG--LALISEAVLRGQALLVNSSQPEPLQ----LHYDKAVSGLRS 104
 DB 1719 AALIEGGQRMAMLSKIQGLQCATRQADIPAIQSSWPKLEQFMKHMNFISGIQY 1778
 OY 105 LTTLLRALG---AKKEAISPDDAASAPLRTIT 134
 DB 1779 LAGLSTLPGNPAVVASMAFS--AALISPLPTST 1809

RESULT 15

O43380 PRELIMINARY; PRT; 480 AA.
 AC O43380;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Putative pp70 ribosomal protein S6 kinase.
 GN ASPK11.
 OS Avena sativa (Oat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Avenae; Avena.
 OC NCBI_TaxID=4498;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. RHIANNON; TISSUE=ALEURONE;
 RX MEDLINE=95284341; PubMed=7766874;
 RA Huttly A.K., Phillips A.L.;
 RT "qJbberellin regulated expression in oat aleurone cells of two kinases
 RL Plant Mol. Biol. 27:1043-1052(1995).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; X79992; CAA56313.1; -.
 DR HSSP; P05132; ICP.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR000961; Pkinase_C.
 DR InterPro: IPR002290; Ser_thr_Pkinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR Pfam; PF00433; Pkinase_C; 1.
 DR Prodom; PD000001; Euk_Pkinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00133; S_TKc_X; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Ribosomal protein;
 KW Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 480 AA; 53532 MW; 33596A10IDBB077D CRC64;

Query Match 9.3%; Score 79; DB 10; Length 480;
 Best Local Similarity 22.4%; Pred. No. 17;

Matches 32; Conservative 26; Mismatches 65; Indels 20; Gaps 4;

OY 30 AEHCSLNENITVPDTKVFNFYAMKREVGQAAVEWOGIALISEA--VLRGQALLVNSSOP 87
 DB 51 AVHTPEASTIVTPDSPA-----PLAEGSEEVYNDPVITKRSHSLVGPILLVSQSLP 102
 OY 88 WEPLQLHYDKAVSGLRSITTTLLRALGAQKEAISPDDAASAPLRTITADTFKRL----- 141
 ||| ::| | | | : :||| : : : | |

DB 103 LSKLLHSESSSDJ--LECLSKRKQSNQALSDDELSTNFKNEAVGLDNFEVLYALVGOG 160
 OY 142 --FRVYSNFLRGKIKLYTGACR 162
 DB 161 AFGKYYQVRMKGTSEIYAMKVMR 183
 :|| ::| :|| : |

Search completed: January 2, 2003, 15:12:54
 Job time : 31.4199 secs

THIS PAGE BLANK (USPTO)

XX

PS Disclosure; Page 181; 8pp; Japanese.

CC All of the Cys residues in the SQ are labelled "SH". Megakaryocyte-platelet growth factor contains human EPO as an active principle. CC Human EPO has a megakaryocyte colony-stimulating activity and increases the ratio of small acetyl cholinesterase positive cell (S&ChE-) which is immature megakaryocyte. Human EPO effects CC megakaryocyte-platelet system other than an erythrocyte growth CC effect. Megakaryocyte-platelet growth is usable as a remedy for CC diseases caused by a platelet decrease.

XX

SQ Sequence 166 AA;

Query Match 100.0%; Score 851; DB 8; Length 166;
Best Local Similarity 100.0%; Pred. No. 4.3e-87;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRYLERLYLEAKFAENITTCGAHCGLNENITVPDKVNFYAKRMEVGQQA 60
DB 1 APPRLICDSRYLERLYLEAKFAENITTCGAHCGLNENITVPDKVNFYAKRMEVGQQA 60
OY 61 VEVMOGLALLSEAVLRGQALLVNSQPEPQLHVDKAVSGLSRLTLTLRALGAKREANIS 120
DB 61 VEVMOGLALLSEAVLRGQALLVNSQPEPQLHVDKAVSGLSRLTLTLRALGAKREANIS 120
OY 121 PPDASAAPLRTITADTFPRKLFYRYSNFLRGKIKLTYGECRCRGR 166
DB 121 PPDASAAPLRTITADTFPRKLFYRYSNFLRGKIKLTYGECRCRGR 166

Db

121 PPDASAAPLRTITADTFPRKLFYRYSNFLRGKIKLTYGECRCRGR 166

RESULT 2

AAR23593 ID AAR23593 standard; Protein; 166 AA.

XX

AAR23593;

XX

20-OCT-1992 (first entry)

XX

Recombinant hematopoietic molecule portion 2.

DE

Erythropoietin; EPO: erythrocytes; IL-3; haematopoiesis.

XX

Homo sapiens.

OS

WO9206116-A.

XX

16-APR-1992.

XX

26-SEP-1991; 91WO-US07053.

XX

28-SEP-1990; 90US-058958.

XX

(ORTH) ORTHO PHARM CORP.

XX

Rosen JI;

XX

WPI; 1992-150819/18.

XX

Recombinant haematopoietic molecules useful in treating

XX

anemia(s) - comprise IL-3 or GM-CSF, EPO, G-CSF, IL-3 or M-CSF

XX

and has early and later myeloid differentiation activity

XX

Disclosure; Page 32; 82pp; English.

XX

This protein sequence given comprises the entire amino acid sequence

XX

of human erythropoietin (EPO). EPO leads to the maturation of

XX

erythrocytes and is therefore designated as a late myeloid

XX

differentiation factor (MDF) within the scope of the invention

XX

hybrid molecules were produced which contain at least a portion of an

XX

early MDF and at least a portion of a late MDF covalently linked. The

XX

EPO sequence given is effective within the scope of the invention in

XX

full or in a truncated version. Amino acids 7-161 act as a

CC

full or in a truncated version. Amino acids 7-161 act as a

CC

Late MDF when recombined with an early MDF eg. IL-3.

CC

These compounds can be used to promote haematopoiesis in a patient.

CC

The bonding of the early and late factors allows a very high conc. of

CC

late MDF at the surface of a cell which the early MDF is bound. It

CC

also allows the early MDF to act more specifically to stimulate only

CC

the desired lineage, thus reducing undesirable effects. These

CC

compounds are useful for treating anemias of various origins eg. renal

CC

failure and AIDS. It is easier to produce and administer one

XX

recombinant molecule rather than two separate molecules.

SQ

Sequence 166 AA;

Query Match 100.0%; Score 851; DB 13; Length 166;
Best Local Similarity 100.0%; Pred. No. 4.3e-87;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRYLERLYLEAKFAENITTCGAHCGLNENITVPDKVNFYAKRMEVGQQA 60
DB 1 APPRLICDSRYLERLYLEAKFAENITTCGAHCGLNENITVPDKVNFYAKRMEVGQQA 60
OY 61 VEVMOGLALLSEAVLRGQALLVNSQPEPQLHVDKAVSGLSRLTLTLRALGAKREANIS 120
DB 61 VEVMOGLALLSEAVLRGQALLVNSQPEPQLHVDKAVSGLSRLTLTLRALGAKREANIS 120
OY 121 PPDASAAPLRTITADTFPRKLFYRYSNFLRGKIKLTYGECRCRGR 166
DB 121 PPDASAAPLRTITADTFPRKLFYRYSNFLRGKIKLTYGECRCRGR 166

Db

121 PPDASAAPLRTITADTFPRKLFYRYSNFLRGKIKLTYGECRCRGR 166

RESULT 3

AAW77780 ID AAW77780 standard; Protein; 166 AA.

XX

AAW77780;

XX

24-NOV-1998 (first entry)

XX

Human EPO receptor agonist polypeptide.

XX

Haematopoietic receptor agonist; erythropoietin receptor agonist;

XX

EPO; human; chimeric protein; stem cell expansion; tumour;

XX

infection; autoimmune disease; haematopoietic disorder; therapy;

XX

dendritic cell.

XX

Homo sapiens.

OS

WO9206116-A.

XX

16-APR-1992.

XX

26-SEP-1991; 91WO-US07053.

XX

28-SEP-1990; 90US-058958.

XX

(ORTH) ORTHO PHARM CORP.

XX

Rosen JI;

XX

WPI; 1992-150819/18.

XX

Recombinant haematopoietic molecules useful in treating

XX

anemia(s) - comprise IL-3 or GM-CSF, EPO, G-CSF, IL-3 or M-CSF

XX

and has early and later myeloid differentiation activity

XX

Disclosure; Page 32; 82pp; English.

XX

This protein sequence given comprises the entire amino acid sequence

XX

of human erythropoietin (EPO). EPO leads to the maturation of

XX

erythrocytes and is therefore designated as a late myeloid

XX

differentiation factor (MDF) within the scope of the invention

XX

hybrid molecules were produced which contain at least a portion of an

XX

early MDF and at least a portion of a late MDF covalently linked. The

XX

EPO sequence given is effective within the scope of the invention in

XX

full or in a truncated version. Amino acids 7-161 act as a

CC R1-L1-R2, R2-L1-R1, R1-R2 or R2-R1, where L is a linker and R1 and
 CC R2 are independently selected from: (a) the human EPO receptor
 CC agonist; (b) a human stem cell factor receptor agonist polypeptide
 CC (see AAW77781); (c) a human flt-3 receptor agonist polypeptide (see
 CC AAW77782); (d) a modified human granulocyte colony stimulating factor
 CC (G-CSF) polypeptide (see AAW77783); (e) modified human interleukin-3
 CC polypeptide (see AAW77784); (f) modified human c-mpl ligand polypeptide
 CC (see AAW77785); and (g) a factor selected from the group consisting of
 CC a CSF, a cytokine, a lymphokine, an interleukin and a haematopoietic
 CC growth factor, provided that at least R1 or R2 is selected from (a),
 CC (b) or (c) as above. The multi-functional chimeric haematopoietic
 CC receptor agonist can be used to stimulate the production of

Query Match 100.0%; Score 851; DB 19; Length 166;
 Best Local Similarity 100.0%; Pred. No. 4.3e-87;

Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRYLERYLEAKAEENITTCGAHCNSLMENTIVPTKYNFYAMKMEVGQQA 60
 DB 1 APPRLICDSRYLERYLEAKAEENITTCGAHCNSLMENTIVPTKYNFYAMKMEVGQQA 60
 OY 61 VEWOGIALISEAVLRGQALLVNSOPWEPLOLHVDAKAVSGLSLTLRLALCAQKEAIS 120
 DB 61 VEWOGIALISEAVLRGQALLVNSOPWEPLOLHVDAKAVSGLSLTLRLALCAQKEAIS 120
 OY 121 PPDAASAAPLRTTADTFKRLFRVYSNFIKLGKLTGTGACRGTDR 166
 DB 121 PPDAASAAPLRTTADTFKRLFRVYSNFIKLGKLTGTGACRGTDR 166

RESULT 4

AAW58404
 ID AAW58404 standard; Protein; 166 AA.

AC AAW58404;

DT 12-OCT-1998 (first entry)

DE Human erythropoietin.

KW Erythropoietin receptor agonist; EPO; human; anaemia;

KW haematopoietic deficiency; red blood cell; erythroid progenitor;

KW bone marrow suppression.

OS Homo sapiens.

PN WO9818926-A1.

PD 07-MAY-1998.

PE 23-OCT-1997; 97WO-US18703.

PR 25-OCT-1996; 96US-0034044.

PA (SEAR) SEARLE & CO G D.

PI Feng Y, McWhorter CA, Summers N;

DR WPI: 1998-272221/24.

PT N-PSDB; AAV31031.

Human erythropoietin receptor agonist polypeptide - used to
 stimulate the production of red blood cells in a patient

Claim 1; Page 93; 11pp; English.

CC A claimed human erythropoietin (EPO) receptor agonist polypeptide
 CC comprises a modified EPO amino acid sequence given in AAW58404,
 CC where (a) optionally 1-6 amino acids from the N-terminus and 1-5
 CC from the C-terminus can be deleted, (b) the N-terminus is joined to
 CC the C-terminus directly or through a linker (see AAW58405-12) capable
 CC of joining the N-terminus to the C-terminus, (c) there are new C-
 CC and N-terminal at any two consecutive amino acids from amino acids

CC 23-24 to 38-39, 40-41 to 41-42, 43-44 to 48-49, 50-51 to 57-58,
 CC 77-78 to 82-83, 84-85 to 88-89, and 108-109 to 131-132, and (d)
 CC optionally the agonist polypeptide is preceded by Met, Ala, or
 CC Met-Ala. 60 Of these circularly permuted EPO receptor agonists
 CC (see AAW58413-72) are claimed. Also claimed are: nucleic acid
 CC molecules (see AAV30971-V31030) encoding novel EPO receptor agonists;
 CC a method of producing an EPO receptor agonist using transformed or
 CC transfected host cells; and methods for stimulating the production
 CC of haematopoietic cells, for selective ex vivo expansion of
 CC erythroid progenitors, and treating patients having a haematopoietic
 CC disorder using the EPO receptor agonists. The EPO receptor
 CC agonists using one or more activities of native EPO and may also
 CC show improved haematopoietic cell-stimulating activity and/or an
 CC improved biological activities associated with native EPO and/or have
 CC improved physical properties such as increased solubility,
 CC stability and refold efficiency.

Query Match 100.0%; Score 851; DB 19; Length 166;
 Best Local Similarity 100.0%; Pred. No. 4.3e-87;

Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRYLERYLEAKAEENITTCGAHCNSLMENTIVPTKYNFYAMKMEVGQQA 60
 DB 1 APPRLICDSRYLERYLEAKAEENITTCGAHCNSLMENTIVPTKYNFYAMKMEVGQQA 60
 OY 61 VEWOGIALISEAVLRGQALLVNSOPWEPLOLHVDAKAVSGLSLTLRLALCAQKEAIS 120
 DB 61 VEWOGIALISEAVLRGQALLVNSOPWEPLOLHVDAKAVSGLSLTLRLALCAQKEAIS 120
 OY 121 PPDAASAAPLRTTADTFKRLFRVYSNFIKLGKLTGTGACRGTDR 166
 DB 121 PPDAASAAPLRTTADTFKRLFRVYSNFIKLGKLTGTGACRGTDR 166

RESULT 5

ABB07030
 ID ABB07030 standard; Protein; 166 AA.

AC ABB07030;

DT 21-JUN-2002 (first entry)

DE Modified erythropoietin related gene protein sequence.

KW Modified erythropoietin; EPO.

OS Unidentified.

PN KRL45802-B1.

PD 01-AUG-1998.

PE 31-MAY-1994; 94KR-0012082.

PR 31-MAY-1994; 94KR-0012082.

PA (GLDS) LG CHEM CO LTD.

PI Kim C, Song Y, Lee T;

DR WPI: 2000-234250/20.

PT N-PSDB; ABL50878.

MODIFIED ERYTHROPOIETIN GENE AND EXPRESSION VECTORS THEREOF -
 Disclosure; Page 14; 15pp; Korean.

CC The present invention describes modified erythropoietin (EPO) genes
 CC and expression vectors comprising the genes. The present sequence
 CC represents a protein sequence from the present invention.

```

XX Sequence 166 AA;
SQ Query Match 100.0%; Score 851; DB 21; Length 166;
Best Local Similarity 100.0%; Pred. No. 4.3e-87;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 APPRLICDSRVLERYLLEKAEENITTCGAHCSSLNENITVPDTKYNFYAMKRMVEYGOQA 60
DB 1 APPRLICDSRVLERYLLEKAEENITTCGAHCSSLNENITVPDTKYNFYAMKRMVEYGOQA 60
QY 61 VEWMOGLALLSEAVLRGQALLVNSSQPMPEPLQHDVKAVSGLSLTLLRALGAQKEAIS 120
DB 61 VEWMOGLALLSEAVLRGQALLVNSSQPMPEPLQHDVKAVSGLSLTLLRALGAQKEAIS 120
QY 121 PPDASAAPLRITADTFPKLFRVYSNPLRGKILKLYTGACRTGDR 166
DB 121 PPDASAAPLRITADTFPKLFRVYSNPLRGKILKLYTGACRTGDR 166

```

```

RESULT 6
AAE02641
ID AAE02641 standard; Protein; 166 AA.
XX
XX
AC AAE02641;
XX

```

```

DT 06-AUG-2001 (first entry)
XX
XX

```

```

DE Human erythropoietin (EPO) mature protein.
XX
XX

```

```

KM Human; erythropoietin; EPO; antianaemic; nephrotrophic; anti-HIV;
KM vaccine; haemostatic; immunoglobulin; Ig; EPO deficient disease;
KM anaemia; renal failure; Human Immunodeficiency Virus; HIV;
KM haematopoietic growth factor.
XX
XX

```

```

OS Homo sapiens.
XX
XX
PN WO200136489-A2.
XX

```

```

PD 25-MAY-2001.
XX

```

```

PF 03-NOV-2000; 2000WO-EP10843.
XX

```

```

PR 12-NOV-1999; 99US-0164855.
XX

```

```

PA (MERE ) MERCK PATENT GMBH.
XX

```

```

PI Hartmann A, Brandt S, Rieke E, Sobel C, Lo K, Way JC, Gillies S;
XX

```

```

DR WPI: 2001-367563/38.
XX

```

```

DR N-PSDB; AAD06893.
XX

```

```

PT Novel modified erythropoietin forms such as fusion proteins, comprising
PT Fc portion of an immunoglobulin molecule and a target molecule having
PT the biological activity of erythropoietin forms -
XX
XX

```

```

PS Example 1; Page 22; 58pp; English.
XX

```

```

CC The present sequence is human erythropoietin (EPO) mature protein. EPO
CC has improved biological activity and an extended serum half life greater
CC than 20 hours. The present invention relates to modified EPO forms such
CC as fusion proteins comprising a Fc portion of an immunoglobulin (Ig)
CC molecule and an EPO molecule (Fc-EPO). The Fc portion is fused covalently
CC through its C-terminus directly or indirectly to the EPO molecule, and
CC where the Fc portion as well as EPO portion may be modified or mutated.
CC The invention also relates to non-fused EPO molecules which have a
CC pattern of cysteines or disulphide bonding which is distinct from human
CC or animal EPO. Pharmaceutical compositions containing EPO are useful in
CC the treatment of EPO deficient diseases such as anaemia, renal failure,
CC HIV infection, blood loss and chronic disease that can be treated with
CC haematopoietic growth factor.
XX
XX

```

```

SQ Sequence 166 AA;

```

```

Query Match 100.0%; Score 851; DB 22; Length 166;
Best Local Similarity 100.0%; Pred. No. 4.3e-87;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 APPRLICDSRVLERYLLEKAEENITTCGAHCSSLNENITVPDTKYNFYAMKRMVEYGOQA 60
DB 1 APPRLICDSRVLERYLLEKAEENITTCGAHCSSLNENITVPDTKYNFYAMKRMVEYGOQA 60
QY 61 VEWMOGLALLSEAVLRGQALLVNSSQPMPEPLQHDVKAVSGLSLTLLRALGAQKEAIS 120
DB 61 VEWMOGLALLSEAVLRGQALLVNSSQPMPEPLQHDVKAVSGLSLTLLRALGAQKEAIS 120
QY 121 PPDASAAPLRITADTFPKLFRVYSNPLRGKILKLYTGACRTGDR 166
DB 121 PPDASAAPLRITADTFPKLFRVYSNPLRGKILKLYTGACRTGDR 166

```

```

RESULT 7
AAB66698
ID AAB66698 standard; Protein; 166 AA.
XX
XX
AC AAB66698;
XX

```

```

DT 06-APR-2001 (first entry)
XX
XX

```

```

DE Human erythropoietin protein #2.
XX
XX

```

```

KM Erythropoietin; EPO; reticulocytes; red blood cell;
KM ethylene glycol; chronic renal failure; AIDS; cancer.
XX
XX

```

```

OS Homo sapiens.
XX
XX

```

```

PN WO200102017-A2.
XX

```

```

PD 11-JAN-2001.
XX

```

```

PF 28-JUN-2000; 2000WO-EP06009.
XX

```

```

PR 02-JUL-1999; 99US-0142243.
XX

```

```

PR 05-AUG-1999; 99US-0147452.
XX

```

```

PR 30-AUG-1999; 99US-0151454.
XX

```

```

PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX

```

```

PI Burg J, Hilger B, Josel H;
XX

```

```

DR WPI: 2001-147051/15.
XX

```

```

PT Novel erythropoietin-glycoprotein conjugate useful for treating
PT diseases correlated with anemia in chronic renal failure patients, AIDS
PT and for treating cancer, is linked to polyethylene glycol through
PT linker -
XX
XX

```

```

PS Claim 19; Fig 2; 40pp; English.
XX

```

```

CC The present invention relates to a conjugate comprising, human
CC erythropoietin glycoprotein (EPO) having at least one free
CC amino group and having in vivo biological activity of causing
CC an increase the production of reticulocytes and red blood cells,
CC covalently linked to 1-3 lower-alkoxy poly(ethylene glycol) groups
CC through a linker. The invention is useful for preparation of
CC medicaments for the treatment of prophylaxis of disease correlated
CC with anemia in chronic renal failure patients (CRF), AIDS and for
CC the treatment of cancer patients undergoing chemotherapy.
XX
XX

```

```

SQ Sequence 166 AA;

```

```

Query Match 100.0%; Score 851; DB 22; Length 166;
Best Local Similarity 100.0%; Pred. No. 4.3e-87;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 APPRLICDSRVLERYLLEKAEENITTCGAHCSSLNENITVPDTKYNFYAMKRMVEYGOQA 60

```

```

Db      1  APPRLICDSRVLEIRYLEEKAENITTTGCAEHCSINENTIVPDTXVNFYAKMRMEVGOA 60
        |||
Oy      61  VEWOGIALISEAVLRGQALVNSQWPPELQIHYDKANVSGIRSLTTLRALGAQKEAIS 120
        |||
Db      61  VEWOGIALISEAVLRGQALVNSQWPPELQIHYDKANVSGIRSLTTLRALGAQKEAIS 120
        |||
Oy      121  PPDASAAPLRTITADTFRKLFPRVYSNFLRGKIKLYTGACRTGDR 166
        |||
Db      121  PPDASAAPLRTITADTFRKLFPRVYSNFLRGKIKLYTGACRTGDR 166
        |||

```

RESULT 8

ABB77897 standard; protein: 166 AA.

ABB77897;

07-OCT-2002 (first entry)

Amino acid sequence of a human erythropoietin (EPO).

Human; erythropoietin; EPO; glycoprotein; reticulocyte production;
red blood cell production; anaemia; chronic renal failure;
acquired immunodeficiency syndrome; AIDS; cancer; bone marrow;
committed erythroid progenitor.

Homo sapiens.

WO200249673-A2.

27-JUN-2002.

08-DEC-2001; 2001WO-EP14434.

20-DEC-2000; 2000EP-0127891.

(HOFF) HOFFMANN LA ROCHE & CO AG F.

Burg J, Engel A, Franze R, Hilger B, Schurig HE, Tischer W;
Wozny M;

WPI: 2002-566640/60.

Novel conjugate of erythropoietin glycoprotein with polyethylene glycol, useful for treating diseases correlated with anaemia in chronic renal failure patients and acquired immunodeficiency syndrome -
Claim 26: Fig 2; 40pp; English.

The present sequence represents a human erythropoietin (EPO) protein. It was used to produce conjugates of the invention. The specification describes a conjugate comprising an EPO glycoprotein having an N-terminal alpha-amino group, chosen from human EPO (hEPO) or its analogues (where hEPO is modified by addition of 1-6 glycosylation sites or a rearrangement of a glycosylation site). The glycoprotein is covalently linked to a poly(ethylene glycol) group. The EPO glycoprotein has in vivo biological activity of causing bone marrow cells to increase production of reticulocytes and red blood cells. The conjugate increased circulating half-life and plasma residence time, decreased clearance, increased clinical activity in vivo, improved potency and stability, when compared to unmodified EPO. The EPO conjugate is useful for preparing medicaments for the treatment and prophylaxis of diseases correlated with anaemia in chronic renal failure patients (CRF), acquired immunodeficiency syndrome (AIDS) and for treating cancer patients undergoing chemotherapy. It is also useful for treating patients by stimulating the division and differentiation of committed erythroid progenitors in the bone marrow.

Sequence 166 AA;

Query Match 100.0%; Score 851; DB 23; length 166;
Best Local Similarity 100.0%; Pred. No. 4, 3e-87;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy      1  APPRLICDSRVLEIRYLEEKAENITTTGCAEHCSINENTIVPDTXVNFYAKMRMEVGOA 60
        |||
Db      1  APPRLICDSRVLEIRYLEEKAENITTTGCAEHCSINENTIVPDTXVNFYAKMRMEVGOA 60
        |||
Oy      61  VEWOGIALISEAVLRGQALVNSQWPPELQIHYDKANVSGIRSLTTLRALGAQKEAIS 120
        |||
Db      61  VEWOGIALISEAVLRGQALVNSQWPPELQIHYDKANVSGIRSLTTLRALGAQKEAIS 120
        |||
Oy      121  PPDASAAPLRTITADTFRKLFPRVYSNFLRGKIKLYTGACRTGDR 166
        |||
Db      121  PPDASAAPLRTITADTFRKLFPRVYSNFLRGKIKLYTGACRTGDR 166
        |||

```

RESULT 9

AAM53062 standard; protein: 166 AA.

AAM53062;

25-MAR-2002 (first entry)

Human erythropoietin (hEPO), 166 residue form.

Human; erythropoietin; EPO; hEPO; haemostatic; red blood cell;
blood disorder; anaemia; chronic renal failure; CRF; AIDS;
acquired immunodeficiency syndrome; cancer chemotherapy; haemostatic;
anti-HIV; antianaemic.

Homo sapiens.

WO200187329-A1.

22-NOV-2001.

08-MAY-2001; 2001WO-EP05187.

15-MAY-2000; 2000EP-0110355.

(HOFF) HOFFMANN LA ROCHE & CO AG F.

Papadimitriou A;

WPI: 2002-082943/11.

Composition useful in the treatment of e.g. AIDS comprises an erythropoietin protein, and a multiple charged inorganic anion in a buffer -
Claim 28: Fig 2; 64pp; English.

The invention relates to liquid pharmaceutical compositions comprising an erythropoietin (EPO) protein, a multiple negatively charged inorganic anion in a buffer which maintains the pH of the solution from 5.5-7.0, and optionally at least one excipient. The erythropoietin used in the composition is preferably human (AAM53061 or AAM53062) a human (AAM53064-AAM53107), or an erythropoietin with the C-terminal addition of a C-terminal fragment of human chorionic gonadotropin (AAM53063). Erythropoietin is a glycoprotein essential for the formation of red blood cells and is therefore useful in the treatment of blood disorders

CC characterised by low or defective red blood cell production. The
CC compositions of the invention can be used in the treatment and prevention
CC of anaemia in chronic renal failure patients (CRF), AIDS (acquired
CC immunodeficiency syndrome), and/or for the treatment of cancer patients
CC undergoing chemotherapy. Unlike prior art erythropoietin compositions,
CC the compositions of the invention do not contain human serum albumin
CC (thereby avoiding the possibility of viral infections and allergic
CC reactions associated with this component), are liquid rather than
CC lyophilisates (and therefore do not need to be reconstituted before
CC administration), and are stable at elevated temperatures such as 25
CC degrees Celsius and even 40 degrees Celsius, and therefore can be stored
CC without refrigeration for prolonged periods without degradation and loss
CC of activity. The present sequence represents the 166 residue form
CC of human erythropoietin which is specifically claimed for use in a
CC composition of the invention.
CC
XX
SQ Sequence 166 AA;
Query Match 100.0%; Score 851; DB 23; Length 166;
Best Local Similarity 100.0%; Pred. No. 4.3e-87;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 APRRLICDSRVLEERYLLEAKAEENITTCGAHCSENENITVPDTRKVNFRYAKRMVEVGQA 60
DB 1 APRRLICDSRVLEERYLLEAKAEENITTCGAHCSENENITVPDTRKVNFRYAKRMVEVGQA 60
QY 61 VEWMOGLALISEAVLRGQALLVNSSQPEWPIQLHVDKAVSGRLSTTLRLALGAQKEAIS 120
DB 61 VEWMOGLALISEAVLRGQALLVNSSQPEWPIQLHVDKAVSGRLSTTLRLALGAQKEAIS 120
QY 121 PPDAASAAPLRTITADTRFKLFRRVYSNPLRGKRLKLYTGEACRTGDR 166
DB 121 PPDAASAAPLRTITADTRFKLFRRVYSNPLRGKRLKLYTGEACRTGDR 166
RESULT 10
AAP50298
ID AAP50298 standard; protein: 167 AA.
XX
AC AAP50298;
XX
DT 01-JAN-1980 (first entry)
XX
DE Human recombinant erythropoietin expressed in *Saccharomyces*
DE *cerevisiae*.
XX
KW Erythropoietin; red blood cell; erythrocyte; anaemia; blood;
KW disorder; ds; *Saccharomyces cerevisiae*.
XX
OS Homo sapiens.
XX
PN W08502610-A.
XX
PD 20-JUN-1985.
XX
PE 11-DEC-1984; 84WO-US02021.
XX
PR 30-NOV-1984; 84US-0675298.
PR 13-DEC-1983; 83US-0561024.
PR 21-FEB-1984; 84US-0582185.
PR 28-SEP-1984; 84US-0655841.
XX
PA (KIRI-) KIRIN-AMGEN INC.
XX
XX WPI: 1985-159229/26.
DR N-PSDB: AAN50345.
XX
XX New polypeptide having properties of erythropoietin - is prepd.
PT by cultivation of transformed eucaryotic or procaryotic host
XX
PS Disclosure; Page 82; 113pp; English.
XX
XX Human erythropoietin encoded by this sequence is essential for red

CC blood cell formation and is used for the diagnosis and treatment of
CC blood disorders such as anaemia. Large amounts of EPO may be obtained
CC using recombinant DNA techniques in contrast to small amounts
CC obtained from plasma and urine. This sequence is expressed in *S.*
CC *cerevisiae*. See also AAN50346-50 and AAP50299-P50301.
XX
SQ Sequence 167 AA;
Query Match 100.0%; Score 851; DB 6; Length 167;
Best Local Similarity 100.0%; Pred. No. 4.3e-87;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 APRRLICDSRVLEERYLLEAKAEENITTCGAHCSENENITVPDTRKVNFRYAKRMVEVGQA 60
DB 2 APRRLICDSRVLEERYLLEAKAEENITTCGAHCSENENITVPDTRKVNFRYAKRMVEVGQA 61
QY 61 VEWMOGLALISEAVLRGQALLVNSSQPEWPIQLHVDKAVSGRLSTTLRLALGAQKEAIS 120
DB 62 VEWMOGLALISEAVLRGQALLVNSSQPEWPIQLHVDKAVSGRLSTTLRLALGAQKEAIS 121
QY 121 PPDAASAAPLRTITADTRFKLFRRVYSNPLRGKRLKLYTGEACRTGDR 166
DB 122 PPDAASAAPLRTITADTRFKLFRRVYSNPLRGKRLKLYTGEACRTGDR 167
RESULT 11
AAP50299
ID AAP50299 standard; protein: 167 AA.
XX
AC AAP50299;
XX
DT 01-JAN-1980 (first entry)
XX
DE Human recombinant erythropoietin expressed in *Escherichia coli*.
DE
XX
KW Erythropoietin; red blood cell; erythrocyte; anaemia; blood;
KW disorder; ds; *Escherichia coli*.
XX
OS Homo sapiens.
XX
PN W08502610-A.
XX
PD 20-JUN-1985.
XX
PE 11-DEC-1984; 84WO-US02021.
XX
PR 30-NOV-1984; 84US-0675298.
PR 13-DEC-1983; 83US-0561024.
PR 21-FEB-1984; 84US-0582185.
PR 28-SEP-1984; 84US-0655841.
XX
PA (KIRI-) KIRIN-AMGEN INC.
XX
XX WPI: 1985-159229/26.
DR N-PSDB: AAN50346.
XX
XX New polypeptide having properties of erythropoietin - is prepd.
PT by cultivation of transformed eucaryotic or procaryotic host
XX
PS Disclosure; Page 72; 113pp; English.
XX
XX Human erythropoietin encoded by this sequence is essential for red
CC blood cell formation and is used for the diagnosis and treatment of
CC blood disorders such as anaemia. Large amounts of EPO may be obtained
CC using recombinant DNA techniques in contrast to small amounts
CC obtained from plasma and urine. This sequence is expressed in *E.*
CC *coli*. See also AAN50345, AAN50347-50 and AAP50298, AAP50300-P50301.
XX
SQ Sequence 167 AA;
Query Match 100.0%; Score 851; DB 6; Length 167;
Best Local Similarity 100.0%; Pred. No. 4.3e-87;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEKYLEAKAEENITTCGAHCSLNENITVPDTRKVNPFYAKRMREVGQA 60
 |||||||
 DB 2 APPRLICDSRVLEKYLEKAEENITTCGAHCSLNENITVPDTRKVNPFYAKRMREVGQA 61
 |||||||
 OY 61 VEWMOGLALSEAVLRGQALLVNSSQPEWPLQHDVDAVSGLSRLTTLRALGAQKEAIS 120
 |||||||
 DB 62 VEWMOGLALSEAVLRGQALLVNSSQPEWPLQHDVDAVSGLSRLTTLRALGAQKEAIS 121
 |||||||
 OY 121 PPDASAAPLRTITADTRFKLFRVYSNPLRGKLTGTGEACRTGDR 166
 |||||||
 DB 122 PPDASAAPLRTITADTRFKLFRVYSNPLRGKLTGTGEACRTGDR 167
 |||||||

RESULT 12

ABR77899 standard; protein: 169 AA.

AC ABR77899;

DT 07-OCT-2002 (first entry)

DE Amino acid sequence of a modified human erythropoietin (EPO).

XX Human: erythropoietin; EPO; glycoprotein; reticulocyte production;

KW red blood cell production; anaemia; chronic renal failure;

KW acquired immunodeficiency syndrome; AIDS; cancer; bone marrow;

XX committed erythroid progenitor.

OS Synthetic.

XX Homo sapiens.

XX Key

FT Cleavage-site 1..3

FT Protein 4..174

FT Protein /note- "EPO protein"

PN WO200249673-A2.

XX 27-JUN-2002.

PD 08-DEC-2001; 2001WO-EP14434.

XX 20-DEC-2000; 2000EP-0127891.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Burg J, Engel A, Franze R, Hilger B, Schurig HE, Tischer W;

PI Wozny M;

XX WPI; 2002-566640/60.

XX Novel conjugate of erythropoietin glycoprotein with polyethylene

PT glycol, useful for treating diseases correlated with anaemia in chronic

PT renal failure patients and acquired immunodeficiency syndrome -

XX Disclosure; Page 39; 40pp: English.

XX The present sequence represents a modified human erythropoietin (EPO)

XX protein. The EPO was extended at the N-terminal by a proteolytic

XX cleavage site. It was used to produce conjugates of the invention. The

XX specification describes a conjugate comprising an EPO glycoprotein having

XX an N-terminal alpha-amino group, chosen from human EPO (hEPO) or its

XX analogues (where hEPO is modified by addition of 1-6 glycosylation sites

XX or a rearrangement of a glycosylation site). The glycoprotein is

XX covalently linked to a poly(ethylene glycol) group. The EPO glycoprotein

XX has in vivo biological activity of causing bone marrow cells to increase

XX production of reticulocytes and red blood cells. The conjugate increased

XX circulating half-life and plasma residence time, decreased clearance,

XX increased clinical activity in vivo, improved potency and stability, when

XX compared to unmodified EPO. The EPO conjugate is useful for preparing

XX medicaments for the treatment and prophylaxis of diseases correlated with

CC anaemia in chronic renal failure patients (CRF), acquired

CC immunodeficiency syndrome (AIDS) and for treating cancer patients

CC undergoing chemotherapy. It is also useful for treating patients by

CC stimulating the division and differentiation of committed erythroid

CC progenitors in the bone marrow.

XX Sequence 169 AA;

SO Query Match 100.0%; Score 851; DB 23; Length 169;

Best Local Similarity 100.0%; Pred. No. 4, 4e-87;

Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEKYLEAKAEENITTCGAHCSLNENITVPDTRKVNPFYAKRMREVGQA 60

DB 4 APPRLICDSRVLEKYLEKAEENITTCGAHCSLNENITVPDTRKVNPFYAKRMREVGQA 63

OY 61 VEWMOGLALSEAVLRGQALLVNSSQPEWPLQHDVDAVSGLSRLTTLRALGAQKEAIS 120

DB 64 VEWMOGLALSEAVLRGQALLVNSSQPEWPLQHDVDAVSGLSRLTTLRALGAQKEAIS 123

OY 121 PPDASAAPLRTITADTRFKLFRVYSNPLRGKLTGTGEACRTGDR 166

DB 124 PPDASAAPLRTITADTRFKLFRVYSNPLRGKLTGTGEACRTGDR 169

RESULT 13

ABR77898 standard; protein: 174 AA.

AC ABR77898;

DT 07-OCT-2002 (first entry)

DE Amino acid sequence of a modified human erythropoietin (EPO).

XX Human: erythropoietin; EPO; glycoprotein; reticulocyte production;

KW red blood cell production; anaemia; chronic renal failure;

KW acquired immunodeficiency syndrome; AIDS; cancer; bone marrow;

XX committed erythroid progenitor.

OS Synthetic.

XX Homo sapiens.

XX Key

FT Cleavage-site 1..8

FT Protein 9..174

FT Protein /note- "EPO protein"

PN WO200249673-A2.

XX 27-JUN-2002.

PD 08-DEC-2001; 2001WO-EP14434.

XX 20-DEC-2000; 2000EP-0127891.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Burg J, Engel A, Franze R, Hilger B, Schurig HE, Tischer W;

PI Wozny M;

XX WPI; 2002-566640/60.

XX Novel conjugate of erythropoietin glycoprotein with polyethylene

PT glycol, useful for treating diseases correlated with anaemia in chronic

PT renal failure patients and acquired immunodeficiency syndrome -

XX Disclosure; Page 38-39; 40pp: English.

XX The present sequence represents a modified human erythropoietin (EPO)

XX protein. The EPO was extended at the N-terminal by a proteolytic

XX cleavage site. It was used to produce conjugates of the invention. The

CC specification describes a conjugate comprising an EPO glycoprotein having
CC an N-terminal alpha-amino group, chosen from human EPO (hEPO) or its
CC analogues (where hEPO is modified by addition of 1-6 glycosylation sites
CC or a rearrangement of a glycosylation site). The glycoprotein is
CC covalently linked to a poly(ethylene glycol) group. The EPO glycoprotein
CC has *in vivo* biological activity of causing bone marrow cells to increase
CC production of reticulocytes and red blood cells. The conjugate increased
CC circulating half-life and plasma residence time, decreased clearance,
CC increased clinical activity *in vivo*, improved potency and stability, when
CC compared to unmodified EPO. The EPO conjugate is useful for preparing
CC medicaments for the treatment and prophylaxis of diseases correlated with
CC anaemia in chronic renal failure patients (CRF), acquired
CC immunodeficiency syndrome (AIDS) and for treating cancer patients
CC undergoing chemotherapy. It is also useful for treating patients by
CC stimulating the division and differentiation of committed erythroid
CC progenitors in the bone marrow.

SO Sequence 174 AA;

Query Match 100.0%; Score 851; DB 23; Length 174;

Best Local Similarity 100.0%; Pred. No. 4,6e-87;

Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLEERYLLEAKEAENITTCGAHCSSLNENITVPDTKVFYAMKRMVEYGOQA 60
DB 9 APPRLICDSRVLEERYLLEAKEAENITTCGAHCSSLNENITVPDTKVFYAMKRMVEYGOQA 68

QY 61 VEVWOGLLLSFAYVLRGQALLVNSSQPEPQLDHYDKAVSGLRSLTTLRALGAQKEAIS 120
DB 69 VEVWOGLLLSFAYVLRGQALLVNSSQPEPQLDHYDKAVSGLRSLTTLRALGAQKEAIS 128

QY 121 PPDASAAPLRTITADTFPKLFRVYSNPLRGKLYTGACRGTGR 166
DB 129 PPDASAAPLRTITADTFPKLFRVYSNPLRGKLYTGACRGTGR 174

RESULT 14
ABB77900
ID ABB77900 standard; protein; 174 AA.

AC ABB77900;

DT 07-OCT-2002 (first entry)

DE Amino acid sequence of a modified human erythropoietin (EPO).

KM Human; erythropoietin; EPO; glycoprotein; reticulocyte production;
KM red blood cell production; anaemia; chronic renal failure;
KM acquired immunodeficiency syndrome; AIDS; cancer; bone marrow;
KM committed erythroid progenitor.

OS Synthetic.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Cleavage-site 1..8 /note= "proteolytic cleavage site"

FT Protein 9..174 /note= "EPO protein"

PN WO200249673-A2.

PD 27-JUN-2002.

PF 08-DEC-2001; 2001WO-EP14434.

PR 20-DEC-2000; 2000EP-0127891.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Burg J, Engel A, Franze R, Hilger B, Schurig HE, Tischner W,
PI Wozny M;

DR WPI; 2002-566640/60.

PT Novel conjugate of erythropoietin glycoprotein with polyethylene
PT glycol, useful for treating diseases correlated with anaemia in chronic
PT renal failure patients and acquired immunodeficiency syndrome -

PS Disclosure; Page 39-40; 40pp; English.

CC The present sequence represents a modified human erythropoietin (EPO)
CC protein. The EPO was extended at the N-terminal by a proteolytic
CC cleavage site. It was used to produce conjugates of the invention. The
CC specification describes a conjugate comprising an EPO glycoprotein having
CC an N-terminal alpha-amino group, chosen from human EPO (hEPO) or its
CC analogues (where hEPO is modified by addition of 1-6 glycosylation sites
CC or a rearrangement of a glycosylation site). The glycoprotein is
CC covalently linked to a poly(ethylene glycol) group. The EPO glycoprotein
CC has *in vivo* biological activity of causing bone marrow cells to increase
CC production of reticulocytes and red blood cells. The conjugate increased
CC circulating half-life and plasma residence time, decreased clearance,
CC increased clinical activity *in vivo*, improved potency and stability, when
CC compared to unmodified EPO. The EPO conjugate is useful for preparing
CC medicaments for the treatment and prophylaxis of diseases correlated with
CC anaemia in chronic renal failure patients (CRF), acquired
CC immunodeficiency syndrome (AIDS) and for treating cancer patients
CC undergoing chemotherapy. It is also useful for treating patients by
CC stimulating the division and differentiation of committed erythroid
CC progenitors in the bone marrow.

SO Sequence 174 AA;

Query Match 100.0%; Score 851; DB 23; Length 174;

Best Local Similarity 100.0%; Pred. No. 4,6e-87;

Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLEERYLLEAKEAENITTCGAHCSSLNENITVPDTKVFYAMKRMVEYGOQA 60
DB 9 APPRLICDSRVLEERYLLEAKEAENITTCGAHCSSLNENITVPDTKVFYAMKRMVEYGOQA 68

QY 61 VEVWOGLLLSFAYVLRGQALLVNSSQPEPQLDHYDKAVSGLRSLTTLRALGAQKEAIS 120
DB 69 VEVWOGLLLSFAYVLRGQALLVNSSQPEPQLDHYDKAVSGLRSLTTLRALGAQKEAIS 128

QY 121 PPDASAAPLRTITADTFPKLFRVYSNPLRGKLYTGACRGTGR 166
DB 129 PPDASAAPLRTITADTFPKLFRVYSNPLRGKLYTGACRGTGR 174

RESULT 15
AAP60599
ID AAP60599 standard; protein; 168 AA.

AC AAP60599;

DT 01-JAN-1980 (first entry)

DE Clone lambda HEPOL16 encoding human erythropoietin.

KM Erythropoietin; lambda HEPOL16; recombinant plasmid vector; anaemia;
KM mammal cell culture; 3T3; CHO; Chinese hamster ovary; ss.

OS Homo sapiens.

PN WO8603520-A.

PD 19-JUN-1986.

PF 03-DEC-1985; 85WO-US02405.

PR 22-JAN-1985; 85US-0693258.

PR 04-DEC-1984; 84US-0677813.

PR 03-JAN-1985; 85US-0688622.

PA (GENE-) GENETICS INST INC.

PA (FRIT/) FRITSCHE E.
 XX
 PI Fritsch E, Hewick RM, Jacobs K;
 XX
 DR WPI; 1986-169459/26.
 DR N-PSDB; AAN60519.
 XX
 PT Prodn. of human cDNA clone expressing erythropoietin - for mass
 XX Prodn. of erythropoietin, useful for treating anaemia
 XX
 PS Disclosure; Page 20; 61pp; English.
 XX
 CC A recombinant plasmid vector expressing this clone is expressed in e.
 CC g 3T3 or CHO cell cultures. The produced erythropoietin is useful
 CC for treatment of anaemia, especially renal anaemia. The cloned gene
 CC expresses high levels of the protein and thus provides a means of
 CC mass production. See also AAN60513-18, AAN60520-21 and AAP60598.
 XX
 SO Sequence 188 AA;
 Query Match 100.0%; Score 851; DB 7; Length 188;
 Best Local Similarity 100.0%; Pred. No. 5.1e-87;
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 APPRLICDSRYLERYLLLEAKHAENTTTCAGHCSLNENITVPPTKVFYAKRMREYGOOA 60
 Db 23 APPRLICDSRYLERYLLLEAKHAENTTTCAGHCSLNENITVPPTKVFYAKRMREYGOOA 82
 QY 61 VEVWOGIALLSEAVLRGOALLVNSQWPEPLQHLVDKAVSGLSRTTLRALGAKKKAIS 120
 Db 83 VEVWOGIALLSEAVLRGOALLVNSQWPEPLQHLVDKAVSGLSRTTLRALGAKKKAIS 142
 QY 121 PPDAAASAPLRTITADTFRKLFRVYSNPLRGKILKLTGECRTGDR 166
 Db 143 PPDAAASAPLRTITADTFRKLFRVYSNPLRGKILKLTGECRTGDR 188

Search completed: January 2, 2003, 15:11:28
 Job time : 33.0967 secs


```

OY      1  APPRLICSRVLERLLLEAKAEANTTGCACHSCLNENTTTPDRKVPNFYAKRREYGOA  60
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1  APPRLICSRVLERLLLEAKAEANTTGCACHSCLNENTTTPDRKVPNFYAKRREYGOA  60
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      61  VEWOGALLSEAVLRGOALLVNSSOPWPEPLQLHVDKAVSGLSRLTTLRLALGAOKEAIS  120
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      61  VEWOGALLSEAVLRGOALLVNSSOPWPEPLQLHVDKAVSGLSRLTTLRLALGAOKEAIS  120
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      121  PPDASAPPLRTITADTFRKLFRVYSNPLRGKLTLYTGEACRTGDR  166
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      121  PPDASAPPLRTITADTFRKLFRVYSNPLRGKLTLYTGEACRTGDR  166
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
US-10-014-363-2
; Sequence 2, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischler, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 166
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-014-363-2

Query Match
Best Local Similarity 100.0%; Score 851; DB 12; Length 166;
Pred. No. 1,1e-84;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDKVNFYAMKRMEVGOQA 60
D 1 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDKVNFYAMKRMEVGOQA 60
QY 61 VEWOGIALISEAVLRGOALLVNSOPWEPLOLHVDAVSGLSITLTLRALGAOKRAIS 120
D 61 VEWOGIALISEAVLRGOALLVNSOPWEPLOLHVDAVSGLSITLTLRALGAOKRAIS 120
QY 121 PPDAASAAPLRTITADTFRKLFRVYSNPLRGKLLKLYGECACRTGDR 166
D 121 PPDAASAAPLRTITADTFRKLFRVYSNPLRGKLLKLYGECACRTGDR 166
```

```

RESULT 3
US-10-014-363-4
; Sequence 4, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischler, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 169
; TYPE: PRN
; ORGANISM: CHO/dhfr-
US-10-014-363-4

Query Match
Best Local Similarity 100.0%; Score 851; DB 12; Length 169;
Pred. No. 1,1e-84;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDKVNFYAMKRMEVGOQA 60
D 4 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDKVNFYAMKRMEVGOQA 63
QY 61 VEWOGIALISEAVLRGOALLVNSOPWEPLOLHVDAVSGLSITLTLRALGAOKRAIS 120
D 64 VEWOGIALISEAVLRGOALLVNSOPWEPLOLHVDAVSGLSITLTLRALGAOKRAIS 123
QY 121 PPDAASAAPLRTITADTFRKLFRVYSNPLRGKLLKLYGECACRTGDR 166
```

```

D 124 PPDAASAAPLRTITADTFRKLFRVYSNPLRGKLLKLYGECACRTGDR 169
```

```

RESULT 4
US-10-014-363-3
; Sequence 3, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischler, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 174
; TYPE: PRN
; ORGANISM: CHO/dhfr-
US-10-014-363-3
```

```

Query Match
Best Local Similarity 100.0%; Score 851; DB 12; Length 174;
Pred. No. 1,2e-84;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDKVNFYAMKRMEVGOQA 60
D 9 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDKVNFYAMKRMEVGOQA 68
QY 61 VEWOGIALISEAVLRGOALLVNSOPWEPLOLHVDAVSGLSITLTLRALGAOKRAIS 120
D 69 VEWOGIALISEAVLRGOALLVNSOPWEPLOLHVDAVSGLSITLTLRALGAOKRAIS 128
QY 121 PPDAASAAPLRTITADTFRKLFRVYSNPLRGKLLKLYGECACRTGDR 166
D 129 PPDAASAAPLRTITADTFRKLFRVYSNPLRGKLLKLYGECACRTGDR 174
```

```

RESULT 5
US-10-014-363-5
; Sequence 5, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischler, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 174
; TYPE: PRN
; ORGANISM: CHO/dhfr-
US-10-014-363-5
```

```

Query Match
Best Local Similarity 100.0%; Score 851; DB 12; Length 174;
Pred. No. 1,2e-84;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF053356.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.93
; OTHER INFORMATION: SWISSPROT HIT: P01588, EVALU8 9.00e-28
; OTHER INFORMATION: EST_HUMAN HIT: AA662379.1, EVALU8 3.00e-10
US-09-864-761-48502

Query Match          34.5%; Score 294; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.8e-25;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 VGGQAVEWQGLALSEAVIRGQALLVNSSQPEPIQLHVDKAVSGIRSLTTLRALGAQ 115
DB 1 VGGQAVEWQGLALSEAVIRGQALLVNSSQPEPIQLHVDKAVSGIRSLTTLRALGAQ 60

RESULT 9
US-09-975-063-2
; Sequence 2, Application US/09975063
; Patent No. US2002004525A1
; GENERAL INFORMATION:
; APPLICANT: POWELL, Jerry S.
; TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE: HIGH LEVEL
; NUMBER OF SEQUENCES: EXPRESSION IN STABLY TRANSFECTED MAMMALIAN CELLS
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dorsey & Whitney LLP
; STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.7/7/8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/975,063
; FILING DATE: 10-Oct-2001
; ATTORNEY/AGENT INFORMATION:
; NAME: Roberts, Mark W.
; REGISTRATION NUMBER: 46,160
; REFERENCE/DOCKET NUMBER: 500582.03 (112893.109)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-903-8728
; TELEFAX: 206-903-8820
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-975-063-2

Query Match          13.0%; Score 111; DB 10; Length 30;
Best Local Similarity 83.3%; Pred. No. 8.4e-06;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 APRRLICDSRYLERLYLLEAKENITTCGA 30
DB 1 APRRLIDSRVLERLYLLEAKEXITTDGCA 30

RESULT 10
US-10-011-858-2
; Sequence 2, Application US/10011858
; Patent No. US20020137145A1
; GENERAL INFORMATION:
; APPLICANT: POWELL, Jerry S.
```

```
; TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE: HIGH LEVEL.
; EXPRESSION IN STABLY TRANSFECTED MAMMALIAN CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dorsey & Whitney LLP
; STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS-WORD
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/011,858
; FILING DATE: 05-No. US20020137145A1-2001
; ATTORNEY/AGENT INFORMATION:
; NAME: Roberts, Mark W.
; REGISTRATION NUMBER: 46,160
; REFERENCE/DOCKET NUMBER: 500582.13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-903-8728
; TELEFAX: 206-903-8820
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-011-858-2

Query Match          13.0%; Score 111; DB 12; Length 30;
Best Local Similarity 83.3%; Pred. No. 8.4e-06;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 APRRLICDSRYLERLYLLEAKENITTCGA 30
DB 1 APRRLIDSRVLERLYLLEAKEXITTDGCA 30

RESULT 11
US-09-919-703-9
; Sequence 9, Application US/09919703
; Patent No. US20020165129A1
; GENERAL INFORMATION:
; APPLICANT: Krystal, Gerald
; APPLICANT: Radkin, Simon W.
; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
; TITLE OF INVENTION: Cell Death
; FILING DATE: 30216/003004
; CURRENT APPLICATION NUMBER: US/09/919,703
; FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 1998-0459
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 60/008,233
; PRIOR FILING DATE: 1995-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO: 9
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-09-919-703-9

Query Match          12.7%; Score 108; DB 9; Length 20;
```

STREET: 87 Cambridge
CITY: Cambridge
STATE: Massachusetts

Page 5

Fri Jan 3 10:09:19 2003

us-09-853-731-2.rapb

Page 6

```

1      ?
2      ?
3      ?
4      ?
5      ?
6      ?
7      ?
8      ?
9      ?
10     ?
11     ?
12     ?
13     ?
14     ?
15     ?
16     ?
17     ?
18     ?
19     ?
20     ?
21     ?
22     ?
23     ?
24     ?
25     ?
26     ?
27     ?
28     ?
29     ?
30     ?
31     ?
32     ?
33     ?
34     ?
35     ?
36     ?
37     ?
38     ?
39     ?
40     ?
41     ?
42     ?
43     ?
44     ?
45     ?
46     ?
47     ?
48     ?
49     ?
50     ?
51     ?
52     ?
53     ?
54     ?
55     ?
56     ?
57     ?
58     ?
59     ?
60     ?
61     ?
62     ?
63     ?
64     ?
65     ?
66     ?
67     ?
68     ?
69     ?
70     ?
71     ?
72     ?
73     ?
74     ?
75     ?
76     ?
77     ?
78     ?
79     ?
80     ?
81     ?
82     ?
83     ?
84     ?
85     ?
86     ?
87     ?
88     ?
89     ?
90     ?
91     ?
92     ?
93     ?
94     ?
95     ?
96     ?
97     ?
98     ?
99     ?
100    ?

```

COUNTRY: USA
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/945,182
FILING DATE: 31-Aug-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/808,324
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME/Agent: Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5202-D
TELECOMMUNICATIONS INFORMATION:
TELEPHONE: 617 876-8260
TELEFAX: 617 876-5651

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Query Match	8.5%;	Score 72;	DB 9;	Length 321;
Best Local Similarity	27.2%;	Fred No. 3.8;		
Matches	31;	Conservative 11;	Mismatches 48;	Indels 24; Gaps 3;
OY	57	GOQAEVWGQIALISEVRLRCQAL	LVNSQPEPQLQHVDAVSGLS	LTTLRLALGAQK 116
		: : :	: : :	: : :
Db	77	GMEVEYDVGQL-----	-RIQPKQCLTELRAMGEL	LAGEAEARARGPOO 120
		: : :	: : :	: : :
OY	117	EAI5PPAA5AAPLITITADTFRK	LFERYXSNFLRKL-----	KLYNGEACRTG 164
			: : :	: : :
Db	121	P--PPPLRSLTGRCGRVRPPQ	ERALLVFIISQRNLEFAEKRE	LDGSAEAAAGG 172

Search completed: January 2, 2003, 15:13:50
Job time : 9.52568 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:10:05 ; Search time 14.0423 Seconds
(without alignments)
1136.446 Million cell updates/sec

Title: US-09-853-731-2

Perfect score: 851

Sequence: 1 APRRLIDSRLEKRYLEAK.....NFLRGKLTLYGECACRTGDR 166

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	851	100.0	193	1 ZOHU	erythropoietin pre
2	769.5	90.4	192	1 J00173	erythropoietin pre
3	764.5	89.8	192	1 I84613	erythropoietin pre
4	718	84.4	188	1 I46083	erythropoietin pre
5	706	83.0	192	1 S28148	erythropoietin pre
6	690.5	81.1	194	1 I46401	erythropoietin pre
7	686	80.6	192	1 A24902	erythropoietin pre
8	685.5	80.6	195	1 JC7699	erythropoietin - r
9	683	80.3	190	2 I46578	erythropoietin - p
10	638	75.0	175	2 I46199	erythropoietin - d
11	90	10.6	353	2 G02729	thrombopoietin - h
12	89	10.5	353	2 I80105	thrombopoietin - h
13	88	10.3	323	2 AB0323	erythropoietin pre
14	87.5	10.3	346	2 AE0959	ribonucleoside-dip
15	86	10.1	286	2 A55530	solute binding rec
16	85	10.0	339	2 A83274	megakaryocyte grow
17	83.5	9.8	296	2 A10443	UDP-N-acetylpyruv
18	80.5	9.5	3033	1 GNMVJ8	probable 2-hydroxy
19	79.5	9.3	3033	1 GNMVJ8	genome polypeptide
20	79	9.3	480	2 T35681	probable sensory h
21	78.5	9.2	813	2 A56526	ribosomal protein
22	78.5	9.2	897	2 A54696	ATP-dependent heli
23	78	9.2	348	2 T35450	EGF receptor subst
24	78	9.2	455	2 H97693	ABC transporter AT
25	78	9.2	455	2 AG2919	methylinine utilis
26	78	9.2	747	1 S36741	conserved hypothet
27	77.5	9.1	242	1 AD1928	probable copper-tr
28	77	9.0	451	2 S75569	hypothetical prote
29	77	9.0	548	2 BA2281	hypothetical prote

ALIGNMENTS

RESULT 1
ZOHU
erythropoietin precursor [validated] - human
C:Species: Homo sapiens (man)
C:Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 08-Dec-2000
C:Accession: A01855; A24744; A22210; S56178
R:Jacobs, K.; Shoemaker, C.; Ruderstorff, R.; Neill, S.D.; Kaufman, R.J.; Mufson, A.;
Nature 313, 806-810, 1985
A>Title: Isolation and characterization of genomic and cDNA clones of human erythro
A:Reference number: A01855; MUID:85137899; PMID:3838366
A:Accession: A01855
A:Molecule type: mRNA; DNA
A:Residues: 1-193 <JAC>
A:Cross-references: GB:X02157; GB:X02158
R:Lin, F.K.; Suggs, S.; Ijin, C.H.; Browne, J.K.; Smalting, R.; Egrle, J.C.; Chen, K.K
Proc. Natl. Acad. Sci. U.S.A. 82, 7580-7584, 1985
A>Title: Cloning and expression of the human erythropoietin gene.
A:Reference number: A24744; MUID:86067948; PMID:3865178
A:Accession: A24744
A:Molecule type: DNA
A:Residues: 1-193 <LIN>
A:Cross-references: GB:M11319; NID:9182197; PIDN:AAA52400.1; PID:9182198
R:Lin, F.H.; Everett, R.; Wang, F.F.; Arakawa, T.; Goldwasser, E.
J. Biol. Chem. 261, 3116-3121, 1986
A>Title: Structural characterization of human erythropoietin.
A:Reference number: A25384; MUID:86140080; PMID:3949763
A:Accession: A25384
A:Molecule type: protein
A:Residues: 28-86, 'Q', 87-193 <LA1>
A:Experimental source: urine
A:Note: forms without the carboxyl-terminal residue and the four carboxyl-terminal re
R:Yanagawa, S.; Hirade, K.; Ohnoka, H.; Sasaki, R.; Chiba, H.; Ueda, M.; Goto, M.
J. Biol. Chem. 259, 2707-2710, 1984
A>Title: Isolation of human erythropoietin with monoclonal antibodies.
A:Reference number: A22210; MUID:84135751; PMID:6698989
A:Accession: A22210
A:Molecule type: protein
A:Residues: 28-29, 'X', 31-33, 'L', 35-50, 'X', 52-53, 'D', 55, 'G', 57 <YAN>
R:Matsumoto, S.; Ikura, K.; Ueda, M.; Sasaki, R.
Plant Mol. Biol. 27, 1163-1172, 1995
A>Title: Characterization of a human glycoprotein (erythropoietin) produced in cultur
A:Reference number: S56178; MUID:95284365; PMID:7766897
A:Accession: S56178
A:Molecule type: protein
A:Residues: 28-33, 'X', 35-37 <MTS>
A:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C:Genetics:
A:Gene: GDB:EPO
A:Cross-references: GDB:119110; OMIM:133170
A:Map position: 7q21.3-7q22.1
A:Introns: 5/1; 53/3; 82/3; 142/3
C:Function:

60 kD chaperonin I
bacterioferritin X
mandelate racemase
groEL protein - Ba
hypothetical prote
conserved hypothet
ITS beta (AF305057
ATP-dependent heli
RF2 protein - salm
thrombopoietin pre
ribonucleoside-dip
chaperonin, 60 kD
chaperonin 60 - Co
PST1 protein - yea
probable transport
hypothetical prote

A:Description: the primary inducer of erythrocyte formation
C:Superfamily: erythropoietin
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-27/Domain: signal sequence #status predicted <Sig>
F:28-193/Product: erythropoietin #status experimental <MAT>
F:34-188,56-60/Disulfide bonds: #status experimental
F:51,65,110/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:152/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match
Best Local Similarity 100.0%; Score 851; DB 1; Length 193;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLELYLEAKAEENITTCGAHCSLNENITVPDTKVFYAKRMVEVGQA 60
DB 28 APPRLICDSRVLELYLEAKAEENITTCGAHCSLNENITVPDTKVFYAKRMVEVGQA 87
OY 61 VEWOGIALISEAVLRGQALLVNSQPEPQLHVDKAVGSLITLLRALGAOKEAIS 120
DB 88 VEWOGIALISEAVLRGQALLVNSQPEPQLHVDKAVGSLITLLRALGAOKEAIS 147
OY 121 PPDASAAPLRTITADTFERKLFYRYSNPLRGKLTLYTGECRCRGDR 166
DB 148 PPDASAAPLRTITADTFERKLFYRYSNPLRGKLTLYTGECRCRGDR 193

RESULT 2
JQ0173
erythropoietin precursor - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C>Date: 07-Sep-1990 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C:Accession: JQ0173
R:Lin, F.K.; Lin, C.H.; Tai, P.H.; Browne, J.K.; Eglye, J.C.; Smalling, R.; Fox, G.M.;
Gene 44, 201-209, 1986
A:Title: Monkey erythropoietin gene: cloning, expression and comparison with the human
A:Reference number: JQ0173; MUID:87055236; PMID:877922
A:Accession: JQ0173
A:Molecule type: mRNA
A:Residues: 1-192 <LIN>
A:Cross-references: GB:M18189; GB:M15818; GB:M15819; GB:M18188; NID:9342093; PIDN:AAA368
A:Experimental source: kidney
C:Comment: This protein is the principal hormone involved in the regulation of erythrocy
C:Function: Erythropoietin is produced by kidney or liver of adult mammals and by liver
A:Superfamily: the primary inducer of erythrocyte formation
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-27/Domain: signal sequence #status predicted <Sig>
F:28-192/Product: erythropoietin #status predicted <MAT>
F:34-187,56-60/Disulfide bonds: #status predicted
F:51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:152/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match
Best Local Similarity 90.4%; Score 769.5; DB 1; Length 192;
Matches 152; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

OY 1 APPRLICDSRVLELYLEAKAEENITTCGAHCSLNENITVPDTKVFYAKRMVEVGQA 60
DB 28 APPRLICDSRVLELYLEAKAEENITTCGAHCSLNENITVPDTKVFYAKRMVEVGQA 87
OY 61 VEWOGIALISEAVLRGQALLVNSQPEPQLHVDKAVGSLITLLRALGAOKEAIS 120
DB 88 VEWOGIALISEAVLRGQALLVNSQPEPQLHVDKAVGSLITLLRALGAOKEAIS 146
OY 121 PPDASAAPLRTITADTFERKLFYRYSNPLRGKLTLYTGECRCRGDR 166
DB 147 PPDASAAPLRTITADTFERKLFYRYSNPLRGKLTLYTGECRCRGDR 192

RESULT 3
184613
erythropoietin precursor - rhesus macaque

C:Species: Macaca mulatta (rhesus macaque)
C>Date: 02-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C:Accession: I84613
R:Men, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A:Title: Erythropoietin structure-function relationships: High degree of sequence hom
A:Reference number: I46083; MUID:93372347; PMID:8364201
A:Accession: I84613
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-192 <RES>
A:Cross-references: GB:L10609; NID:9342095; PIDN:AAA36842.1; PID:9342096
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C:Function: Erythropoietin is produced by kidney or liver of adult mammals and by live
A:Superfamily: the primary inducer of erythrocyte formation
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-27/Domain: signal sequence #status predicted <Sig>
F:28-187/Product: erythropoietin #status predicted <MAT>
F:34-187,56-60/Disulfide bonds: #status predicted
F:51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:152/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match
Best Local Similarity 89.8%; Score 764.5; DB 1; Length 192;
Matches 150; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

OY 1 APPRLICDSRVLELYLEAKAEENITTCGAHCSLNENITVPDTKVFYAKRMVEVGQA 60
DB 28 APPRLICDSRVLELYLEAKAEENITTCGAHCSLNENITVPDTKVFYAKRMVEVGQA 87
OY 61 VEWOGIALISEAVLRGQALLVNSQPEPQLHVDKAVGSLITLLRALGAOKEAIS 120
DB 88 VEWOGIALISEAVLRGQALLVNSQPEPQLHVDKAVGSLITLLRALGAOKEAIS 146
OY 121 PPDASAAPLRTITADTFERKLFYRYSNPLRGKLTLYTGECRCRGDR 166
DB 147 PPDASAAPLRTITADTFERKLFYRYSNPLRGKLTLYTGECRCRGDR 192

RESULT 4
I46083
erythropoietin precursor - cat (fragment)
C:Species: Felis silvestris catus (domestic cat)
C>Date: 16-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C:Accession: I46083
R:Men, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A:Title: Erythropoietin structure-function relationships: High degree of sequence hom
A:Reference number: I46083; MUID:93372347; PMID:8364201
A:Accession: I46083
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-188 <MEN>
A:Cross-references: GB:L10606; NID:9163820; PIDN:AAA30607.1; PID:9163821
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C:Function: Erythropoietin is produced by kidney or liver of adult mammals and by live
A:Superfamily: the primary inducer of erythrocyte formation
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-27/Domain: signal sequence (fragment) #status predicted <Sig>
F:23-188/Product: erythropoietin #status predicted <MAT>
F:29-183,51-55/Disulfide bonds: #status predicted
F:46,60,105/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:148/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match
Best Local Similarity 84.4%; Score 718; DB 1; Length 188;
Matches 140; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLELYLEAKAEENITTCGAHCSLNENITVPDTKVFYAKRMVEVGQA 60
DB 23 APPRLICDSRVLELYLEAKAEENITTCGAHCSLNENITVPDTKVFYAKRMVEVGQA 82

OY 61 VEWMOGLALLSEAVLRGQALLVNSSQPWEPIQLHVDKAVSGLSRLTTLRALGAQKEAIS 120
 |||||||
 Db 83 VEWMOGLALLSEAVLRGQALLVNSSQPSETIQLHVDKAVSSIRSLTSLRALGAQKEAIS 142
 |||||||

OY 121 PPDASAAPLRTITADTFRKLEFRVYSNPLRGKLTLYTGECACRTGDR 166
 |||||||
 Db 143 LPEATSAAPLRTITADTFRKLEFRVYSNPLRGKLTLYTGECACRTGDR 188
 |||||||

RESULT 5

S28148
 erythropoietin precursor - rat

C:Species: Rattus norvegicus (Norway rat)
 C:Date: 22-Nov-1993 #sequence_revision 15-Nov-1996 #text_change 21-Jul-2000
 C:Accession: S28148; 162743
 R:Nagao, M.; Suga, H.; Okano, M.; Masuda, S.; Narita, H.; Ikura, K.; Sasaki, R.
 Biochim. Biophys. Acta 1171, 99-102, 1992
 A:Title: Nucleotide sequence of rat erythropoietin.
 A:Reference number: S28148; MUID:93042015; PMID:1420369
 A:Accession: S28148
 A:Molecule type: mRNA
 A:Residues: 1-192 <NAG>
 A:Cross-references: GB:D10763; NID:g220735; PIDN:BAA01593.1; PID:g220736
 R:Men, D.; Boissel, J.
 Blood 82, 1507-1516, 1993
 A:Title: Erythropoietin structure-function relationships: High degree of sequence homol
 A:Reference number: 146083; MUID:9337347; PMID:8364201
 A:Accession: 162743
 A:Status: translated from GB/EMBL/DBD
 A:Molecule type: mRNA
 A:Residues: 4-192 <RES>
 A:Cross-references: GB:L10608; NID:g204060; PIDN:AAA1126.1; PID:g204061
 C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by liver c
 C:Function:
 A:Description: the primary inducer of erythrocyte formation
 C:Superfamily: erythropoietin
 C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-192/Product: erythropoietin #status predicted <MAT>
 F:33-187,55-165/Disulfide bonds: #status predicted
 F:50,64,109/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.0%; Score 706; DB 1; Length 192;
 Best Local Similarity 82.5%; Pred. No. 7.8e-61;
 Matches 137; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLEAKEENITTCGAECSCSLNENTVDPDKVNFYAMKRMVEYGOOA 60
 |||||||
 Db 27 APPRLICDSRVLEERYLLEAKEENITTCGAECSCSLNENTVDPDKVNFYAMKRMVEYGOOA 86
 |||||||

OY 61 VEWMOGLALLSEAVLRGQALLVNSSQPWEPIQLHVDKAVSGLSRLTTLRALGAQKEAIS 120
 |||||||
 Db 87 VEWMOGLALLSEAVLRGQALLVNSSQPESQLHIDKAIISGRSLTSLRALGAQKEAIS 146
 |||||||

OY 121 PPDASAAPLRTITADTFRKLEFRVYSNPLRGKLTLYTGECACRTGDR 166
 |||||||
 Db 147 PPDATQAAPLRTITADTFRKLEFRVYSNPLRGKLTLYTGECACRTGDR 192
 |||||||

RESULT 6

146401
 erythropoietin precursor - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 16-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
 C:Accession: 146401; 147077
 R:Fu, P.; Evans, B.; Lim, G.B.; Moritz, K.; Wintour, E.M.
 Mol. Cell. Endocrinol. 93, 107-116, 1993
 A:Title: The sheep erythropoietin gene: molecular cloning and effect of hemorrhage on p
 A:Reference number: 146401; MUID:93351736; PMID:8349021
 A:Accession: 146401
 A:Status: translated from GB/EMBL/DBD
 A:Molecule type: mRNA

A:Residues: 1-194 <FOX>
 A:Cross-references: EMBL:L24681; NID:g395049; PIDN:CAA80848.1; PID:g395050
 R:Men, D.; Boissel, J.
 Blood 82, 1507-1516, 1993
 A:Title: Erythropoietin structure-function relationships: High degree of sequence hom
 A:Reference number: 146083; MUID:93372347; PMID:8364201
 A:Accession: 147077
 A:Status: translated from GB/EMBL/DBD

A:Molecule type: mRNA
 A:Residues: 4-15, 'V', '17-107, 'P', '109-194 <MEN>
 A:Cross-references: GB:L10610; NID:g165876; PIDN:AAA31518.1; PID:g165877
 C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
 C:Function:
 A:Description: the primary inducer of erythrocyte formation
 C:Superfamily: erythropoietin
 C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-194/Product: erythropoietin #status predicted <MAT>
 F:34-189,56-60/Disulfide bonds: #status predicted
 F:51,65,110/Binding site: carbohydrate (Ser) (covalent) #status predicted
 F:154/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 81.1%; Score 690.5; DB 1; Length 194;
 Best Local Similarity 82.0%; Pred. No. 2.5e-59;
 Matches 137; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

OY 1 APPRLICDSRVLEERYLLEAKEENITTCGAECSCSLNENTVDPDKVNFYAMKRMVEYGOOA 60
 |||||||
 Db 28 APPRLICDSRVLEERYLLEAKEENITTCGAECSCSLNENTVDPDKVNFYAMKRMVEYGOOA 87
 |||||||

OY 61 VEWMOGLALLSEAVLRGQALLVNSSQPWEPIQLHVDKAVSGLSRLTTLRALGAQKEAIS 120
 |||||||
 Db 88 VEWMOGLALLSEAVLRGQALLVNSSQPCEALRLHVDKAVSGLSRLTTLRALGAQKEAIS 147
 |||||||

OY 121 PPDASAAPLRTITADTFRKLEFRVYSNPLRGKLTLYTGECACRTGDR 166
 |||||||
 Db 148 LPEATSAAPLRTITADTFRKLEFRVYSNPLRGKLTLYTGECACRTGDR 194
 |||||||

RESULT 7

A24902
 erythropoietin precursor - mouse

C:Species: Mus musculus (house mouse)
 C:Date: 25-Oct-1987 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
 C:Accession: A24902; A24901
 R:Shoemaker, C.B.; Mitsch, L.D.
 Mol. Cell. Biol. 6, 849-858, 1986
 A:Title: Murine erythropoietin gene: cloning, expression, and human gene homology.
 A:Reference number: A24902; MUID:87039105; PMID:3773894
 A:Accession: A24902
 A:Molecule type: DNA
 A:Residues: 1-192 <SHO>
 A:Note: the authors translated the codon TTA for residue 12 as Phe, TTA for residue 4
 R:McDonald, J.D.; Lin, F.K.; Goldwasser, E.
 Mol. Cell. Biol. 6, 842-848, 1986
 A:Title: Cloning, sequencing, and evolutionary analysis of the mouse erythropoietin g
 A:Reference number: A24901; MUID:87039104; PMID:3022133
 A:Accession: A24901
 A:Molecule type: DNA
 A:Residues: 1-67, 'P', '69-192 <MCD>
 A:Cross-references: GB:M12930; NID:g193086; PIDN:AAA37570.1; PID:g387152
 C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
 C:Genetics:
 A:Introns: 5/1; 52/3; 81/3; 141/3
 C:Function:
 A:Description: the primary inducer of erythrocyte formation
 C:Superfamily: erythropoietin
 C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-192/Product: erythropoietin #status predicted <MAT>
 F:33-187,55-165/Disulfide bonds: #status predicted
 F:50,64,109/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 80.6%; Score 686; DB 1; Length 192;
Best Local Similarity 79.5%; Pred. No. 6,7e-59;
Matches 132; Conservative 14; Mismatches 20; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLLEAKENITTCGAHCNSLNTITVPDTKVNFYAKRMVEVGQA 60
DB 27 APPRLICDSRVLEERYLLLEAKENITTCGAHCNSLNTITVPDTKVNFYAKRMVEVGQA 86
OY 61 VEWMOGLALISEAVLRGQALLVNSSOPPEPQLQHVDAVSGRLSTLTLLRALGAKKEAIS 120
DB 87 LEVMOGLALISEAVLRGQALLVNSSOPPEPQLQHVDAVSGRLSTLTLLRALGAKKEAIS 146
OY 121 PDDAASAPLRTITADTFKRLFRVYSNPLRGKLTGTGACRRTGDR 166
DB 147 PDDTPPAPLRTITADTFKRLFRVYSNPLRGKLTGTGACRRTGDR 192

RESULT 8

JC7699

erythropoietin - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001

C:Accession: JC7699

R:Vialta, A.; Wu, D.; Margalith, M.; Hobart, P.

Biochem. Biophys. Res. Commun. 284, 823-827, 2001

A:Title: Rabbit Epo gene and cDNA: Expression of rabbit Epo after intramuscular injectio

A:Reference number: JC7699; MUID:21290682; PMID:11396976

A:Contents: kidney

A:Accession: JC7699

A:Molecule type: DNA

A:Residues: 1-195 <VID>

A:Cross-references: GB:AF290943

C:Comment: This protein, a heavily glycosylated 34k protein produced in the fetal liver

cytes.

C:Genetics:

A:Gene: epo

C:Superfamily: erythropoietin

C:Keywords: glycoprotein; kidney

Query Match 80.6%; Score 685.5; DB 2; Length 195;
Best Local Similarity 81.4%; Pred. No. 7,7e-59;
Matches 136; Conservative 12; Mismatches 18; Indels 1; Gaps 1;

OY 1 APPRLICDSRVLEERYLLLEAKENITTCGAHCNSLNTITVPDTKVNFYAKRMVEVGQA 60
DB 29 APPRLICDSRVLEERYLLLEAKENITTCGAHCNSLNTITVPDTKVNFYAKRMVEVGQA 88
OY 61 VEWMOGLALISEAVLRGQALLVNSSOPPEPQLQHVDAVSGRLSTLTLLRALGAKKEAIS 120
DB 89 VEWMOGLALISEAVLRGQALLVNSSOPPEPQLQHVDAVSGRLSTLTLLRALGAKKEAIS 148
OY 121 PDDAASAPLRTITADTFKRLFRVYSNPLRGKLTGTGACRRTGDR 166
DB 149 PDDAASAPLRTITADTFKRLFRVYSNPLRGKLTGTGACRRTGDR 195

RESULT 9

I46578

erythropoietin - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999

C:Accession: I46578

R:Wen, D.; Boissel, J.

Blood 82, 1507-1516, 1993

A:Title: Erythropoietin structure-function relationships: High degree of sequence homolo

A:Reference number: I46083; MUID:93372347; PMID:8364201

A:Accession: I46578

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-190 <MEN>

A:Cross-references: GB:L10607; NID:g164445; PIDN:AAA31029.1; PID:g164446

C:Superfamily: erythropoietin

Query Match 80.3%; Score 683; DB 2; Length 190;
Best Local Similarity 82.1%; Pred. No. 1,3e-58;
Matches 138; Conservative 7; Mismatches 21; Indels 2; Gaps 1;

OY 1 APPRLICDSRVLEERYLLLEAKENITTCGAHCNSLNTITVPDTKVNFYAKRMVEVGQA 60
DB 23 APPRLICDSRVLEERYLLLEAKENITTCGAHCNSLNTITVPDTKVNFYAKRMVEVGQA 82
OY 61 VEWMOGLALISEAVLRGQALLVNSSOPPEPQLQHVDAVSGRLSTLTLLRALGAKKEAIS 120
DB 83 LEVMOGLALISEAVLRGQALLVNSSOPPEPQLQHVDAVSGRLSTLTLLRALGAKKEAIS 142
OY 121 PDDAASAPLRTITADTFKRLFRVYSNPLRGKLTGTGACRRTGDR 166
DB 143 PDDAASAPLRTITADTFKRLFRVYSNPLRGKLTGTGACRRTGDR 190

RESULT 10

I46199

erythropoietin - dog (fragment)

C:Species: Canis lupus familiaris (dog)

C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 23-Jul-1999

C:Accession: I46199

R:Wen, D.; Boissel, J.

Blood 82, 1507-1516, 1993

A:Title: Erythropoietin structure-function relationships: High degree of sequence hom

A:Reference number: I46083; MUID:93372347; PMID:8364201

A:Accession: I46199

A:Molecule type: mRNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-175 <MEN>

A:Cross-references: GB:I13027; NID:g290087; PIDN:AAA30842.1; PID:g552347

C:Superfamily: erythropoietin

Query Match 75.0%; Score 638; DB 2; Length 175;
Best Local Similarity 81.0%; Pred. No. 2,6e-54;
Matches 124; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLLEAKENITTCGAHCNSLNTITVPDTKVNFYAKRMVEVGQA 60
DB 23 APPRLICDSRVLEERYLLLEAKENITTCGAHCNSLNTITVPDTKVNFYAKRMVEVGQA 82
OY 61 VEWMOGLALISEAVLRGQALLVNSSOPPEPQLQHVDAVSGRLSTLTLLRALGAKKEAIS 120
DB 83 LEVMOGLALISEAVLRGQALLVNSSOPPEPQLQHVDAVSGRLSTLTLLRALGAKKEAIS 142
OY 121 PDDAASAPLRTITADTFKRLFRVYSNPLRGKLTGTGACRRTGDR 163
DB 143 PDDAASAPLRTITADTFKRLFRVYSNPLRGKLTGTGACRRTGDR 175

RESULT 11

G02729

thrombopoietin - human

C:Species: Homo sapiens (man)

C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 05-Nov-1999

C:Accession: G02729

R:Im, S.

Submitted to the EMBL Data Library, May 1996

A:Reference number: H01637

A:Accession: G02729

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-353 <IXA>

A:Cross-references: EMBL:U59493; NID:g1401245; PIDN:AAB03392.1; PID:g1401246

C:Genetics:

A:Gene: htpo

Query Match 10.6%; Score 90; DB 2; Length 353;
Best Local Similarity 26.3%; Pred. No. 0.61;
Matches 41; Conservative 20; Mismatches 75; Indels 20; Gaps 5;

OY 1 APPRLICDSRVLEERYLLLEAKENITTCGAHCNSLNTITVPDTKVNFYAKRMVEVGQA 60

```

Db      24 APP--ACDLRVLSKLRSDSHVLSKLSQCEVHPRLPTPVLLPAVDFSLGEMKWTQMEETKA 81
      61 VEWMOGLALISEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGLSRLTTLRALGAQKEA 118
      Db      82 QDILGAVTLLLEGVMAARQGLPFTCLSSLLGOLSGVRLILGALQSL-----LGTQ--- 132
      QY      119 ISPPDASAPLRTTADTFRKLFRRVYSNFLRGK 154
      Db      133 -LPPQG-----RTTAHKDPNMIPLSFQHLRGKVR 161

RESULT 12
180105
thrombopoietin precursor - human
N:Alternate names: c-MPL ligand; megakaryocyte growth and development factor precursor
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 20-Jun-2000
C:Accession: I59281; 180105; S45331; S48740; I38672; I52610
R:Poster, D.C.; Sprecher, C.A.; Grant, F.J.; Kramer, J.M.; Kuijper, J.L.; Holly, R.D.; W
proc. Natl. Acad. Sci. U.S.A. 91, 13023-13027, 1994
A:Title: Human thrombopoietin: gene structure, cDNA sequence, expression, and chromosome
A:Reference number: I59281; MUID:95108091; PMID:7809166
A:Accession: I59281
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-353 <RE2>
A:Cross-references: GB:I36051; NID:9533214; PIDN:AAC37568.1; PID:9533215
A:Accession: I80105
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-353 <RES>
A:Cross-references: GB:I36052; NID:9533216; PIDN:AAC37566.1; PID:9533217
R:de Sauvage, F.J.; Hass, P.E.; Spencer, S.D.; Malloy, B.E.; Gurney, A.L.; Spencer, S.A.
D.V.; Eaton, D.L.
Nature 369, 533-538, 1994
A:Title: Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl ligand.
A:Reference number: S45331; MUID:94261202; PMID:8202154
A:Accession: S45331
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-353 <SAU>
A:Cross-references: GB:I33410; NID:9506826; PIDN:AAA9857.1; PID:9506827
R:Solima, Y.; Akahori, H.; Seki, N.; Horii, T.; Ogami, K.; Kato, T.; Shimada, Y.; Kawamura
FEBS Lett. 353, 57-61, 1994
A:Title: Molecular cloning and chromosomal localization of the human thrombopoietin gene
A:Reference number: S48740; MUID:95010765; PMID:7926023
A:Accession: S48740
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <SOH>
A:Cross-references: GB:D32046; NID:9577319; PIDN:BAAO6807.1; PID:9577320
R:Barclay, T.D.; Bogenberger, J.; Hunt, P.; Li, Y.S.; Lu, H.S.; Martin, F.; Chang, M.S.;
Cell 77, 1117-1124, 1994
A:Title: Identification and cloning of a megakaryocyte growth and development factor tha
A:Reference number: A54463; MUID:94291201; PMID:8020099
A:Accession: I38672
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-112, 'E', 114-353 <RE3>
A:Cross-references: EMBL:U1025; NID:9511223; PIDN:AAA0553.1; PID:958078
R:Gurney, A.L.; Khang, W.J.; Xie, M.H.; Malloy, B.E.; Eaton, D.L.; de Sauvage, F.J.
Blood 85, 981-988, 1995
A:Title: Genomic structure, chromosomal localization, and conserved alternative splice f
A:Reference number: I52610; MUID:95152076; PMID:7849319
A:Accession: I52610
A:Status: preliminary
A:Molecule type: DNA
A:Cross-references: GB:S76771; NID:9914225; PIDN:AA33390.1; PID:9914226
C:Genetics:
A:Gene: GDB:THPO; MGD
A:Cross-references: GDB:374007; OMIM:600044

```

```

A:Map position: 3q26.3-3q27
A:Introns: 5/1: 47/3: 76/3: 132/3
C:Keywords: alternative splicing; cytokine; glycoprotein
Query Match      10.5%; Score 89; DB 2; Length 353;
Best Local Similarity 26.3%; Pred. No. 0.77;
Matches 41; Conservative 20; Mismatches 75; Indels 20; Gaps 5;

      QY      1 APRILCSRVLEERYLLEKKEENITTCGEHCSLNENITVPDTKYNFYAKRMEYGOA 60
      Db      24 APP--ACDLRVLSKLRSDSHVLSKLSQCEVHPRLPTPVLLPAVDFSLGEMKWTQMEETKA 81
      QY      61 VEWMOGLALISEAVL--RGQALLVNSSQPWEPLQLHVDKAVSGLSRLTTLRALGAQKEA 118
      Db      82 QDILGAVTLLLEGVMAARQGLPFTCLSSLLGOLSGVRLILGALQSL-----LGTQ--- 132
      QY      119 ISPPDASAPLRTTADTFRKLFRRVYSNFLRGK 154
      Db      133 -LPPQG-----RTTAHKDPNMIPLSFQHLRGKVR 161

RESULT 13
AB0323
ribonucleoside-diphosphate reductase (EC 1.17.4.1) beta chain [imported] - Yersinia p
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C:Accession: AB0323
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tlball, R.W.; Holden, M.T.G.; Prentice, M.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barr
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0323
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-323 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92889.1; PID:915980631; GSPDB:GN00175
C:Genetics:
A:Gene: nrdf
C:Superfamily: ribonucleoside-diphosphate reductase beta
C:Keywords: oxidoreductase
Query Match      10.3%; Score 88; DB 2; Length 323;
Best Local Similarity 25.2%; Pred. No. 0.86;
Matches 34; Conservative 20; Mismatches 59; Indels 22; Gaps 5;

      QY      38 NITVPDTKYNFYAKRMEYGOA VEWMOGLALISEAVLRGQALLVNSSQPWEPLQLHVD- 96
      Db      2 NVVKKPITRISALNMWKE-DKQDLEVMN--RLTSNFWLPEKVPKLSNDIPSWATLTPHEQ 58
      QY      97 ---KAVSGLSRLTTLRALGAQ---KEAISPDPASAPLRTTADTFRKLFRRVYSNFLR 150
      Db      59 LTRVFTGTLTDITONTGAPALIKDAITPBEAIFSNISFMENHARSYSSTL-- 116
      QY      151 GKLKLYTGACRTGD 165
      Db      117 -----CLTSD 121

RESULT 14
AE0959
Solute binding receptor protein [imported] - Salmonella enterica subsp. enterica sero
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: This species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AE0959
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Parr
, S.; Woude, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se

```

Fri Jan 3 10:09:20 2003

us-09-853-731-2.rpr

Page 6

A:Reference number: AB0502; PMID:11677608
A:Accession: AE0959
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-346 <PAR>
A:CROSS-references: GB:AL513382; PIDN:CAMD03169.1; PID:J16504804; GSFDB:GN00176
A:Genetics:
A:Gene: STR3952

[illegible]

```

RESULT 15
A:Accession: A55530
M: megakaryocyte growth and development factor, long form - human
N: Alternate names: MPL ligand, long form
C: Species: Homo sapiens (man)
C: Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text-change 07-May-1999
C: Accession: A55530
R: Chang, M.; McNinch, J.; Basu, R.; Shutter, J.; Hsu, R.; Perkins, C.; Mar, V.; Suggs,
J. Biol. Chem. 270: 511-514, 1995
A: Title: Cloning and characterization of the human megakaryocyte growth and development
A: Reference number: A55530; MUID:95122483; PMID:7822271
A: Accession: A55530
A: Status: Preliminary: not compared with conceptual translation
A: Molecule type: DNA
A: Residues: 1-286 <CHA>
A: Cross-references: GB:U17071
C: Genetics:
A: Gene: MGDF
A: Map position: 3q26.3
A: Keywords: alternative splicing; cytokine

```

Query Match	10.1%	Score 86;	DB 2;	Length 286;
Best Local Similarity	26.6%;	Pred No.1.2;	Mismatches 41;	Conservative 18;
Matches	41;	Conservative 18;	Mismatches 41;	Indels 20;
				Gaps 5;
QY	1	APRLICRSRYERLLLEAKENENTTTGCAEHCSINENTIVPDTVFNFYAMKRMVEGOOA	60	
Db	24	APR--ACDLRYLSKILRSHSVHLSRSCQPEPHPLPTVLLPVPDFSLGEMWTQMEETKA	81	
QY	61	VEWOGGLALLSRAVL--RGQALLVNSQSPWEFLQIHDYKAVSGLSITLTLLRALGAQKEA	118	
Db	82	QDILGAVTLLEGGVAAAGQ.GPTCLSSLTQLSGQVYKLLGALQSL-----LGTQ---	132	
QY	119	ISPPDASAPLRTTTADTFRLLRFVFNSTLRGK	152	
Db	133	LPPDG-----RTTAHKDPNAITFLSFQHLRKG	159	

Search completed: January 2, 2003, 15:13:27
Job time : 15.0423 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:07:09 ; Search time 8.52568 Seconds
(without alignments)
807.569 Million cell updates/sec

Title: US-09-853-731-2

Perfect score: 851
Sequence: 1 APRRLIDSRVLEKRYLLEAK.....NFLRGKIKLYTGECRTGDR 166

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	851	100.0	193	EPO_HUMAN	P01588 homo sapien
2	769.5	90.4	192	EPO_MACFA	P07865 macaca fasc
3	764.5	89.8	192	EPO_MACMU	O28513 macaca mula
4	711	83.5	192	EPO_PELCA	P33708 felis silve
5	706	83.0	192	EPO_RAT	P26676 rattus norv
6	697.5	82.0	192	EPO_BOVIN	P48617 bos taurus
7	694	81.6	192	EPO_MOUSE	P07321 mus musculu
8	690.5	81.1	194	EPO_SHEEP	P33709 ovis aries
9	683	80.3	190	EPO_PIG	P49157 sus scrofa
10	638	75.0	175	EPO_CANFA	P33707 canis famli
11	109	12.8	352	TPO_CANFA	P42705 canis famli
12	89	10.5	353	TPO_HUMAN	P40225 homo sapien
13	82	9.6	548	CH60_BUCAP	O51832 buchnera ap
14	81	9.5	551	CH60_BUCAP	O59177 buchnera ap
15	80.5	9.5	3033	POLG_HCVJ8	P26661 h genome po
16	79.5	9.3	552	CH60_PSEST	O33500 pseudomonas
17	78.5	9.2	897	EP15_MOUSE	P42567 mus musculu
18	78	9.2	747	ATCS_SYNP7	P37279 synchococc
19	77	9.0	548	CH60_BUCAP	P25750 buchnera ap
20	76.5	9.0	386	CTBP_DROME	O46036 drosophila
21	75.5	8.9	543	CH60_BARBA	P35635 bartonella
22	75	8.8	547	CH60_LEGPN	P28878 legionella
23	74.5	8.8	809	HRPB_ECOLI	P37024 escherichia
24	74	8.7	326	TPO_RAT	P47745 rattus norv
25	74	8.7	544	CH60_FRATU	P94798 francisella
26	74	8.7	552	CH60_COXBU	P19421 coxiella bu
27	74	8.7	1089	IMB3_YEAST	P33337 saccharomyc
28	74	8.7	1564	PDRA_YEAST	P51533 saccharomyc
29	73.5	8.6	830	GCL2_MOUSE	O91hd2 mus musculu
30	73.5	8.6	837	GCL2_HUMAN	O92830 homo sapien
31	73	8.6	263	YH25_DEIRA	O83030 deinococcus
32	72.5	8.5	762	SLAP_ACEKI	P22258 acetogenium
33	72	8.5	356	TPO_MOUSE	P40226 mus musculu

34	71	8.3	547	1	CH60_PASMU	O59687 pasteurella
35	70.5	8.3	217	1	YMBD_ECOLI	P76909 escherichia
36	70.5	8.3	381	1	MODD_MYCAN	O48919 mycobacteri
37	70.5	8.3	896	1	EP15_HUMAN	P42566 homo sapien
38	70	8.2	319	1	R1R4_SALTY	P17424 salmonella
39	70	8.2	544	1	CH60_AERSA	O68309 aeromonas s
40	69.5	8.2	544	1	CH60_METCO	P29842 neisseria g
41	69.5	8.2	545	1	CH60_PSEBU	P48216 pseudomonas
42	69.5	8.2	547	1	CH60_PSEAE	P30718 pseudomonas
43	69.5	8.2	551	1	CH60_AMORS	P26004 amoeba prot
44	69.5	8.2	907	1	GACS_PSEST	P48027 pseudomonas
45	69.5	8.2	1327	1	TNKL_HUMAN	O95271 homo sapien

ALIGNMENTS

RESULT 1
EPO_HUMAN STANDARD: PRT: 193 AA.
AC P01588; Q9UHA0; Q9UE25; Q9UD20;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor (Epoetin).
GN EPO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Carnivora; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85137899; PubMed=3838366;
RA Jacobs K., Shoemaker C., Rudersdorf R., Neill S.D., Kaufman R.J.,
RA Mufson A., Seehra J., Jones S.S., Hewick R., Fritsch E.F.,
RA Kawakita M., Shimizu T., Miyake T.;
RT "Isolation and characterization of genomic and cDNA clones of human
RT erythropoietin.";
RL Nature 313:806-810(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067948; PubMed=3865178;
RA Lin F.-K., Sugars S., Lin C.-H., Browne J.K., Smalling R., Egrie J.C.,
RA Chen K.K., Fox G.M., Martin F., Stabinsky Z., Badrawi S.M., Lai P.-H.,
RA Goldwasser E.;
RT "Cloning and expression of the human erythropoietin gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7580-7584(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99018118; PubMed=9799793;
RA Gloeckner G., Scherer S., Schattevoy R., Boright A., Weber J.,
RA Tsui L.-C., Rosenthal A.;
RT "Large-scale sequencing of two regions in human chromosome 7q22:
RT analysis of 650 kb of genomic sequence around the EPO and CUGT1 loci
RT reveals 17 genes.";
RL Genome Res. 8:1060-1073(1998).
RN [4]
RP SEQUENCE FROM N.A.
RX Rupert J.L., Hochachka P.W.;
RT "Erythropoietin gene sequence in the Quechua, a high altitude native
RT population.";
RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
RN [5]
RP SEQUENCE OF 58-193 FROM N.A., AND VARIANTS HEPATOCELLULAR CARCINOMA.
RX MEDLINE=93384593; PubMed=8396923;
RA Funakoshi A., Muta H., Baba T., Shimizu S.;
RT "Gene expression of mutant erythropoietin in hepatocellular
RT carcinoma.";
RL Biochem. Biophys. Res. Commun. 195:717-722(1993).
RN [6]
RP SEQUENCE OF 28-193, AND DISULFIDE BONDS.
RX TISSUE-Urine;
MEDLINE=86140080; PubMed=3949763;


```

DE Erythropoietin precursor.
GN EPO.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP
RC SEQUENCE FROM N.A.
RX MEDLINE=93372347; PubMed=8364201;
RA Men D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA Czelusniak J., Goodman M., Bunn H.F.;
RT "Erythropoietin structure-function relationships: high degree of
RT sequence homology among mammals.";
RL Blood 82:1507-1516(1993).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASSES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; L10609; AAA36842.1; -.
DR HSSP; P01588; ICN4.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythropo.
DR Pfam; PF00758; EPO_TPO; 1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO; 1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 27
FT CHAIN 28 192
FT DISULFID 34 187
FT DISULFID 56 60
FT CARBOHYD 51 51
FT CARBOHYD 65 65
FT CARBOHYD 110 110
FT CARBOHYD 152 152
SQ SEQUENCE 192 AA; 21081 MW; 275560A264628CD1 CRC64;

Query Match 89.8%; Score 764.5; DB 1; Length 192;
Best Local Similarity 90.4%; Pred. No. 2.6e-68;
Matches 150; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

QY 1 APRRLCDRYLERYLLEAKEAENITTCGAECSSLENITVPDTKYNFYAMKMEVGQDA 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 28 APRPVGDSRYLERLLEAKEAENITTCGSCSSCLSENITVPDTKYNFYAMKRIEVGQDA 87
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 VEVWQGLALSEAVLRGALLVNSQPEWPIQLQHVDAKAVSGLRSLTTLRALGAKKAIS 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 88 VEVWQGLALSEAVLRGQAVLANSSQPEPIQLHMDKAISGIRSLITTLRALGAG--EAI 146
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 PPDAAAPLRTITADPFRKLPRVYSNLRGKRLKLTGEACHTGR 166
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 147 LPDAAAPLRTITADPFRKLFRIYSNLRGKRLKLTGEACRGR 192

RESULT 4
EPO_FELICA STANDARD: PRT: 192 AA.
AC P33706;
JT 01-FEB-1994 (Rel. 28, Created)
OT 01-OCT-1996 (Rel. 34, Last sequence update)

```

DT	16-OCT-2001 (Rel. 40, last annotation update)
DE	Erythropoietin precursor.
GN	Epo.
OS	Felis silvestris catus (Cat).
OC	Eukaryota; Eukarya; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX	NCBI_taxid=9685;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE-Kidney;
RA	Goodman R.E., Bell R.G.;
RL	Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RP	[2]
RX	SEQUENCE OF 5-192 FROM N.A.
RA	MEDLINE=93372347; Pubmed=8364201;
RA	Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA	Ceclunsiak J., Goodman M., Bunn H.P.;
FT	"Erythropoietin structure-function relationships: high degree of
FT	sequence homology among mammals."
RL	Blood 82:1507-1516(1993).
CC	-1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC	REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC	PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASSES.
CC	-1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC	AND BY LIVER OF FETAL OR NONMATAL MAMMALS.
CC	-1- SIMILARITY: BELONGS TO THE Epo / TPO FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: U00685; AAA18282.1; .
DR	EMBL: U0606; AAA30807.1; .
DR	HSSP: P01598; ICMA.
DR	InterPro: IPR001823; EPO_TPO.
DR	InterPro: IPR003013; Erythropo.
DR	Pfam: PF00755;EPO_TPO: 1.
DR	PRINTS: PR00572; ERYTHROPIN.
DR	PROSITE: PS00817; EPO_TPO; 1.
KV	Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT	CHAIN SIGNAL 27
FT	DISULFID 33
FT	DISULFID 187
FT	DISULFID 55
FT	CARBOHYD 50
FT	CARBOHYD 64
FT	CARBOHYD 109
FT	CONFLICT 44
SO	SEQUENCE 192 AA; 2014 MW; 61C5DA0F5E337293 CRC64;
	BY SIMILARITY.
	ERYTHROPOIETIN.
	BY SIMILARITY.
	N-LINKED (GLCNAC. . .) (POTENTIAL).
	N-LINKED (GLCNAC. . .) (POTENTIAL).
	N-LINKED (GLCNAC. . .) (POTENTIAL).
	G > E (IN REF. 2).

Query Match	83.5%	Score 711	DB 1	Length 192
Best Local Similarity	83.7%	Pred. No. 4	96-65	
Matches 139	Conservative	9	Mismatches 18	Indels 0
			Gaps 0	
QY	1 APPRLICDSVLEFRYLLEKAELENTTTCGAHECSINENTIVPDTKVNFYAMKHEVGGQA 60			
Db	27 APPRLICDSVLEFRYLITGAKENENTTTCGACGCSSENTIVPDTKVNFYAMKHEVGGQA 86			
QY	61 VEWMOGLALISFVYRGQALLVNSQPEPILQHVDAVSGLSITTLTALGAKREAS 120			
Db	87 VEWMOGLALISFVYRGQALLVNSQPEPILQHVDAVSGLSITTLTALGAKREAS 146			
QY	121 PPAAASAPLRTITADTERKLPFRYSNPLRGKLLYGEACRTGDR 166			
Db	147 LPFATSAPLRTFVDTLCKLFRLISNPLRGKLLYGEACRTGDR 192			

```

EPO_RAT
EPO_ID EPO_RAT STANDARD: PRT: 192 AA.
AC P29676; P70504;
AD 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_Taxid=10116;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Miscar; TISSUE=Kidney;
RC MEDLINE=93042015; Pubmed=1420369;
RA Nagao M., Suga H., Okano M., Masuda S., Narita H., Ikura K.,
RT Sasaki R.;
RT Nucleotide sequence of rat erythropoietin.";
RT Biochim. Biophys. Acta 1171:99-102(1992).
[2]
RN SEQUENCE OF 4-192 FROM N.A.
RP STRAIN=Sprague-Dawley; TISSUE=Kidney;
RC MEDLINE=93372347; Pubmed=8364201;
RA Wen D., Bolissel J.F.R., Tracy T.E., Mulcahy L.S., Czelusniak J.,
RT Goodman M., Bunn H.F.;
RT Erythropoietin structure-function relationships: high degree of
RT sequence homology among mammals.";
RT Blood 82:1507-1516(1993).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NONMAMMAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

[illegible]

Db 87 VEWQGISLJSLSEALLOAQALOANSQSPPESIQHLHDKAISGLRSLSLTSLRVLGQAKEMLS 146

Oy 121 PPDAASAPLRTITADTFRKLFRLVRYSNFLRGKKLKLYTGACRTGD R 166
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 147 PPDATOAAPLRTITADTFCKLFRVYSNPLRGKKLKLYTGACRRGRD R 192

RESULT 6
EPO_BOVIN STANDARD: PRF: 192 AA.

AC PA4617;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Ox NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BORAN: TISSUE-Kidney;
RX MEDLINE=96257233; PubMed=8666286;
RA Suliman H.B., Majlwa P.A.O., Feldman B.F., Mertens B.,
Logan-Henfrey L.L.;
RT "Cloning of a cDNA encoding bovine erythropoietin and analysis of its
transcription in selected tissues";
RL Gene 171:275-280(1996).

-I FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -I SUBCELLULAR LOCATION: Secreted.
CC -I TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -I SIMILARITY: BELONGS TO THE EPO / TPQ FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on ways
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

CC -----
DR EMBL; LA1354; AAAA41268.1; -.
DR EMBL; U44762; AAA8653.1; -.
DR HSSP; P01588; ICN4.
DR InterPro; IPR001323; EPO.TPQ.
DR InterPro; IPR003013; Erythroptn.
DR Pfam; PF00758; EPO_TPQ_1.
DR PRINTS; PR00272; ERYTHROPNT.
DR PROSITE; PS00817; EPO_TPQ_1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 32 192 ERYTHROPOIETIN.
FT DISULFD 54 58 BY SIMILARTY.
FT DISULFD 54 58 BY SIMILARTY.
FT CARBOHD 49 49 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHD 63 63 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 108 108 N-LINKED (GLCNAG. . .) (POTENTIAL).
SQ SEQUENCE 192 AA; 21075 MW; DBC419022EF7B483A CRC64;

Query Match 82.0%; Score 697.5; DB 1; Length 192:
Best Local Similarity 83.2%; Pred. No. 1.le-61;
Matches 139; Conservative 8; Mismatches 19; Indels 1; Gaps 1;

Oy 1 APPRTCDSRVLERLLTAKEAEENTTTCAECSCISNENTNYDPDVKNFYAKRMVEVGOA 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26 APARLICDSRVLERIITLAAREKENATMGCAESCFSFNENTNYVDIVVFAYAKRMVEVOOA 85
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Oy 61 VEWMGIALLSELAVLRGAQLLVNSQPWPEPIQLAHVDKAVSGIRSLSTTLRALGAOKEAIS 120

[illegible]

```

DR EMBL: AF312033; AAK28825.1; -.
DR EMBL: U11971; CAA72707.1; -.
DR PIR: A24901; A24901.
DR PIR: A24902; A24902.
DR HSP: P01588; ICN4.
DR MG: M61; 95407; EPO.
DR InterPro: IPR001323; EPO_TPO.
DR InterPro: IPR003013; Erythropo.
DR Pfam: PF00758; EPO_TPO; 1.
DR PRINTS: PR00272; ERYTHROPTN.
DR PROSITE: PS00817; EPO_TPO; 1.
FM Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT CHAIN 1 26
FT DISULFID 27 192
FT CARBOHYD 33 187
FT CARBOHYD 50 50
FT CARBOHYD 64 64
FT CARBOHYD 109 109
SO SEQUENCE 192 AA; 21365 MW; 65F94E214EDDEF2E CRC64;

Query Match 81.6%; Score 694; DB 1; Length 192;
Best Local Similarity 80.1%; Pred. No. 2.3e-61;
Matches 133; Conservative 14; Mismatches 19; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLEAKAEANTTGCAGHCSLMENTITVPTKYNFYAKRMREVGQA 60
DB 27 APPRLICDSRVLEERYLLEAKAEANTTGCAGHCSLMENTITVPTKYNFYAKRMREVEBA 86
OY 61 VEWOGIALLSEAVLRGQALLVNSOPWEPLQHLVDKAVSGRSITLLRALGAKKEAIS 120
DB 87 LEVWOGIALLSEAVLRGQALLVNSOPWEPLQHLVDKAVSGRSITLLRALGAKKEAIS 146
OY 121 PPDAA-SAAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECRCRGR 166
DB 147 PPDTPPAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECRCRGR 192

RESULT 8
EPO_SHEEP
ID EPO_SHEEP STANDARD; PRT: 194 AA.
AC P33709; Q28572;
DT 01-FEB-1994 (Rel. 28; Created)
DT 01-FEB-1994 (Rel. 28; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RC MEDLINE=93351736; PubMed=8349021;
RA Fu P., Evans B., Lim G.B., Moritz K., Wintour M.E.;
RT "The sheep erythropoietin gene: molecular cloning and effect of
RT hemorrhage on plasma erythropoietin and renal/liver messenger RNA in
RT adult sheep.";
RL Mol. Cell. Endocrinol. 93:107-116(1993).
RN [2]
RP SEQUENCE OF 4-194 FROM N.A.
RC TISSUE-Kidney;
RC MEDLINE=93372347; PubMed=8364201;
RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA Czelusniak J., Goodman M., Bunn H.F.;
RT "Erythropoietin structure-function relationships: high degree of
RT sequence homology among mammals.";
RL Blood 82:1507-1516(1993).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- SUBCELLULAR LOCATION: Secreted.

```

```

CC CC
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: Z24681; CAA80848.1; -.
DR EMBL: U10610; AAA31518.1; -.
DR HSP: P01588; ICN4.
DR InterPro: IPR001323; EPO_TPO.
DR InterPro: IPR003013; Erythropo.
DR Pfam: PF00758; EPO_TPO; 1.
DR PRINTS: PR00272; ERYTHROPTN.
DR PROSITE: PS00817; EPO_TPO; 1.
FM Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT CHAIN 1 27
FT DISULFID 28 194
FT DISULFID 34 189
FT CARBOHYD 56 60
FT CARBOHYD 51 51
FT CARBOHYD 65 65
FT CARBOHYD 110 110
FT CARBOHYD 16 16
FT CONFLICT 108 108
FT CONFLICT 108 108
SO SEQUENCE 194 AA; 21335 MW; C025AAB0528131A9 CRC64;

Query Match 81.1%; Score 690.5; DB 1; Length 194;
Best Local Similarity 82.0%; Pred. No. 5.3e-61;
Matches 137; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

OY 1 APPRLICDSRVLEERYLLEAKAEANTTGCAGHCSLMENTITVPTKYNFYAKRMREVGQA 60
DB 28 APPRLICDSRVLEERYLLEAKAEANTTGCAGHCSLMENTITVPTKYNFYAKRMREVEBA 87
OY 61 VEWOGIALLSEAVLRGQALLVNSOPWEPLQHLVDKAVSGRSITLLRALGAKKEAIS 120
DB 88 LEVWOGIALLSEAVLRGQALLVNSOPWEPLQHLVDKAVSGRSITLLRALGAKKEAIS 147
OY 121 PPDAA-SAAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECRCRGR 166
DB 148 LPDTPPAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECRCRGR 194

RESULT 9
EPO_PIG
ID EPO_PIG STANDARD; PRT: 190 AA.
AC P49157;
DT 01-FEB-1996 (Rel. 33; Created)
DT 01-FEB-1996 (Rel. 33; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Erythropoietin precursor (Fragment).
GN EPO.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sub.
OX NCBI_TaxID=9623;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RC MEDLINE=93372347; PubMed=8364201;
RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA Czelusniak J., Goodman M., Bunn H.F.;
RT "Erythropoietin structure-function relationships: high degree of
RT sequence homology among mammals.";
RL Blood 82:1507-1516(1993).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A

```

```
CC      PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC      AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC      -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: L10607; AAA31029.1; -.
DR      HSSP: P01588; 1CN4.
DR      InterPro: IPR001323; EPO_TPO.
DR      Pfam: PF00758; EPO_TPO; 1.
DR      PROSITE: PS00817; EPO_TPO; 1.
KW      Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT      NON_TER 1
FT      SIGNAL 1
FT      CHAIN 23 190
FT      DISULFID 29 185
FT      DISULFID 51 55
FT      CARBOHYD 46 46
FT      CARBOHYD 60 60
FT      CARBOHYD 105 105
FT      CARBOHYD 168 168
SQ      SEQUENCE 190 AA; 20888 MW; A75BD6CCE5077E2A CRC64;

Query Match      80.3%; Score 683; DB 1; Length 190;
Best Local Similarity 82.1%; Pred. No. 2.8e-60;
Matches 138; Conservative 7; Mismatches 21; Indels 2; Gaps 1;

OY      1 APPRLICDSRVLEERYLLLEAKEAENITTTGCAEHCSLNEINTVPTKYNFYAMKREMEYGOQA 60
DB      23 APPRLICDSRVLEERYLLLEAKEAENITTTGCAEHCSLNEINTVPTKYNFYAMKREMEYGOQA 82
OY      61 VEWOGIALISEAVNLGQALLVNSQPEWPELQIHDVKAVSGLSLTTLRALGAQKEAIS 120
DB      83 MEWOGIALISEAVNLGQALLVNSQPEWPELQIHDVKAVSGLSLTTLRALGAQKEAIP 142
OY      121 PPDA--ASAAPLRTITADTFKRLFRVYSNPLRGKLTLYTGEACRRTDR 166
DB      143 LPDASPSATPLRTFAVDTLCKLFRNYSNPLRGKLTLYTGEACRRDR 190

RESULT 10
EPO_CANFA
ID      EPO_CANFA      STANDARD;      PRT;      175 AA.
AC      P33707;
DT      01-FEB-1994 (Rel. 28, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Erythropoietin precursor (Fragment).
GN      EPO.
OS      Canis familiaris (Dog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxID=9615;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE-93372347; PubMed-8364201;
RA      Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA      Czelusniak J., Goodman M., Bunn H.F.;
RT      "Erythropoietin structure-function relationships: high degree of
RT      sequence homology among mammals.";
RL      Blood 82:1507-1516(1993).
CC      -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC      REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC      PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC      -1- SUBCELLULAR LOCATION: Secreted.
```

```
CC      -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC      AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC      -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: L13037; AAA30842.1; -.
DR      HSSP: P01588; 1CN4.
DR      InterPro: IPR001323; EPO_TPO.
DR      Pfam: PF00758; EPO_TPO; 1.
DR      PROSITE: PS00817; EPO_TPO; 1.
KW      Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT      NON_TER 1
FT      SIGNAL 1
FT      CHAIN 23 >175
FT      DISULFID 29 >175
FT      DISULFID 51 55
FT      CARBOHYD 46 46
FT      CARBOHYD 60 60
FT      CARBOHYD 105 105
FT      NON_TER 175
SQ      SEQUENCE 175 AA; 19193 MW; B504F8DE86676BF4 CRC64;

Query Match      75.0%; Score 638; DB 1; Length 175;
Best Local Similarity 81.0%; Pred. No. 7e-56;
Matches 124; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

OY      1 APPRLICDSRVLEERYLLLEAKEAENITTTGCAEHCSLNEINTVPTKYNFYAMKREMEYGOQA 60
DB      23 APPRLICDSRVLEERYLLLEAKEAENITTTGCAEHCSLNEINTVPTKYNFYAMKREMEYGOQA 82
OY      61 VEWOGIALISEAVNLGQALLVNSQPEWPELQIHDVKAVSGLSLTTLRALGAQKEAIS 120
DB      83 LEWOGIALISEAVNLGQALLVNSQPEWPELQIHDVKAVSGLSLTTLRALGAQKEAMS 142
OY      121 PPDAASAPLRTITADTFKRLFRVYSNPLRGKLT 153
DB      143 LPDEASAPLRTITADTFKRLFRVYSNPLRGKLT 175

RESULT 11
TPO_CANFA
ID      TPO_CANFA      STANDARD;      PRT;      352 AA.
AC      P42705;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
DE      (C-MPL ligand) (ML) (Megakaryocyte growth and development factor)
GN      (MGDF).
OS      Canis familiaris (Dog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxID=9615;
RN      [1]
RP      SEQUENCE FROM N.A., AND SEQUENCE OF 24-44.
RX      TISSUE-Kidney;
RA      MEDLINE-94291201; PubMed-8020099;
RA      Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.-S., Martin F.,
RA      Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
RA      Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Merewether L.A.,
RA      Closson C., Hsu E., Hukom M.M., Hornkohl A., Choi E., Pangelinan M.,
RA      Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C.,
RA      Shutter J., Chute N., Basu R., Selander L., Trollinger D., Siew L.,
RA      Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
RA      Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,
```

RA Pacific R., Ponting L., Saris C., Wen D., Yung Y.P., Lin H.,
 RA Bosselman R.A.;
 RT "Identification and cloning of a megakaryocyte growth and development
 RL factor that is a ligand for the cytokine receptor Mpl.";
 Cell 77:1117-1117(1994).
 CC -1- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
 CC AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
 CC CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
 CC MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
 CC TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
 CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 DR Interfero: IP001323; EPO_TPO; 1.
 DR Interfero: IP003978; THROMBOPTN.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PRINTS: PS00817; EPO_TPO; 1.
 KM Cysteine, Glycoprotein; Hormone; Signal.
 FT SIGNAL 1 23
 FT CHAIN 1 23
 FT DISULFID 24 352 THROMBOPOIETIN.
 FT DISULFID 26 172 POTENTIAL.
 FT CARBOHYD 135 106 POTENTIAL.
 FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 352 AA; 57641 MW; 024F3B41B061FB08 CRC64;
 Query Match 12.8%; Score 109; DB 1; Length 352;
 Best Local Similarity 24.08; Pred. No. 0.0026;
 Matches 40; Conservative 23; Mismatches 62; Indels 42; Gaps 5;
 QY 1 APPRLICRSYLERILFAKAEINITTCGACBHCNLNENITVPDKRVFVAMKREYVOQA 60
 DB 24 APP--ACDPRLLNKMIDSHVLSRLSCDPIYPLSPVLLPAVDFSLGEMTKOKEOTKA 81
 QY 61 VEVNGLALISEAVL--RGOALLNSQRPWEPDLOLHVRAVSGIRSTLTTLRALGAQKEA 118
 DB 82 QDVWGAVALLDGVLAARGD-----GPSCLISLTQLSGQVRL 120
 QY 119 I-----SPDDAASAPRLTTADPFRKLFVYSNFKRLK 154
 DB 121 LLAGLGLLTGTLPPG-----RTTTHDRPMALFSLFOOLLKGVK 161
 RESULT 12
 TPO_HUMAN
 ID TPO_HUMAN STANDARD; PRT: 353 AA.
 AC P40225; Q13020; Q15790; Q15791; Q15792;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
 DE (Myl) (Megakaryocyte leukemia virus oncogene ligand) (Cmpl ligand)
 DE (Myl) (Megakaryocyte growth and development factor) (MGDF).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX TISSUE=Small Liver;
 RX MEDLINE=94291202; PubMed=8202154;
 RA Spencer S.A., Bardon W.C., Hanzel W.J., Wong S.C., Kuang W.-J.L.,
 RA Oles R.J., Hildgren B., Solberg L.A., Jr., Goeddel D.V., Eaton W.-J.L.;
 RT "Identification of megakaryocyte growth and thrombopoiesis by the c-Mpl
 ligand.";

RL Nature 369:533-538(1994).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Fetal Liver;
 RX MEDLINE=94291201; PubMed=8020009;
 RA Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,
 RA Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
 RA Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Merewether L.A.,
 RA Clogson C., Hsu E., Hockm M., Horikohl A., Choi E., Pangellian M.,
 RA Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C.,
 RA Shutter J., Chute H., Basu R., Selander L., Trollinger D., Siew L.,
 RA Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
 RA Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,
 RA Pacific R., Ponting L., Saris C., Wen D., Yung Y.P., Lin H.,
 RA Bosselman R.A.;
 RT "Identification and cloning of a megakaryocyte growth and development
 RL factor that is a ligand for the cytokine receptor Mpl.";
 Cell 77:1117-1124(1994).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=95108091; PubMed=7809166;
 RA Foster D.C., Sprecher C.A., Grant F.J., Kramer J.M., Kulper J.L.,
 RA Holly R.D., Whitmore T.E., Heipel M.D., Bell L.A.N., Ching A.F.,
 RA McGrane V., Hart C., O'Hara P.J., Lok S.;
 RT "Human thrombopoietin: gene structure, cDNA sequence, expression, and
 RL chromosomal localization.";
 Proc. Natl. Acad. Sci. U.S.A. 91:13023-13027(1994).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=95010765; PubMed=7926023;
 RA Sohma Y., Akahori H., Seki N., Hori T.-A., Ogami K., Kawamura K.,
 RA Miyazaki H.;
 RT "Molecular cloning and chromosomal localization of the human
 RL thrombopoietin gene.";
 FEBS Lett. 353:57-61(1994).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=9512076; PubMed=7849319;
 RA Gurney A.L., Kuang W.-J., Xie M.-H., Malloy B.E., Eaton D.L.,
 RA "de Sauvage F.J.";
 RT "Genomic structure, chromosomal localization, and conserved
 RL alternative splice forms of thrombopoietin.";
 Blood 85:981-988(1995).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX TISSUE=Liver;
 RX MEDLINE=96015174; PubMed=8537317;
 RA Kato T., Ogami K., Shimada Y., Iwamatsu A., Sohma Y., Akahori H.,
 RA Horie K., Okubo A., Kudo Y., Maeda E., Kobayashi K., Ohashi H.,
 RA Ozawa T., Ito H., Kawamura K., Miyazaki H.;
 RT "Purification, characterization of thrombopoietin.";
 J. Biochem. 118:229-236(1995).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX TISSUE=Placenta;
 RX MEDLINE=95122483; PubMed=7822271;
 RA Chang M., McNich J., Basu R., Shutter J., Hsu R., Perkins C., Mar V.,
 RA Suggs S., McNich J., Li L., Lu H., Bartley T., Hunt P., Martin F.,
 RA Samal B., Bogenberger J., Hsu R., Selander L., Trollinger D., Siew L.,
 RA "Cloning and characterization of the human megakaryocyte growth and
 RT development factor (MGDF) gene.";
 J. Biol. Chem. 270:511-514(1995).
 RN [8]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
 RA Im S.H., Lee W.S., Chung K.H.;
 RT "Cloning and sequencing of human thrombopoietin.";
 Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
 CC AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
 CC CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
 CC MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2/TPO-2 AND

DB	133	LPFG-----RTTAHKDPNAIFLSTFQHLKRGV	161
RESULT 13			
ID	CH60_BUCMP	STANDARD:	PRT: 548 AA.
AC	CH60_BUCMP	051832:	
DT	30-MAY-2000	(Rel. 39, Created)	
DT	30-MAY-2000	(Rel. 39, Last sequence update)	
DT	16-OCT-2001	(Rel. 40, Last annotation update)	
DE	60 kDa chaperonin (Protein Cpn60) (groEL protein) (Symbionin).		
GN	GROEL OR MOXA OR GROEL OR SYML.		
OS	Buchnera aphidicola (subsp. Myzus persicae) (Myzus persicae primary endosymbiont).		
OC	Bacteria; Proteobacteria; gamma subdivision; Buchnera.		
OX	NCBI_TaxID=96795;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=98080423; PubMed=9420234;		
RA	Hogenhout S.A., van der Wilk F., Verbeek M., Goldbach R.W.,		
RT	van den Heuvel J.F.J.M.;		
RT	"Potato leafroll virus binds to the equatorial domain of the aphid endosymbiotic GroEL homolog.";		
RL	J. Virol. 72:358-365(1998).		
CC	-1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS CONDITIONS (BY SIMILARITY).		
CC	-1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF 7 SUBUNITS (BY SIMILARITY).		
CC	-1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/sib.ch).		
CC	or send an email to license@sib-sib.ch .		
CC	-----		
DR	EMBL: AF003957; AAC04237.1; -		
DR	HSSP; P06139; IJON.		
DR	InterPro; IPR001844; Chaprinin_Cpn60.		
DR	InterPro; IPR002423; Cpn60_TCP-1.		
DR	Pfam: PF00118; Cpn60_TCP1; 1.		
DR	PRINTS; PR00298; CHAPERONIN60.		
DR	PRINTS; PR00304; TCOMPLEXTCP1.		
DR	PROSITE; PS00296; CHAPERONIN5_CPN60; 1.		
KW	Chaperone; ATP-binding.		
SEQ	SEQUENCE 548 AA; 57883 MW; AD61917965E4184A CRC64;		
Query Match	9.6%; Score 82; DB 1; Length 548;		
Best Local Similarity	24.2%; Pred. NO. 2.1;		
Matches	43; Conservative 33; Mismatches 70; Indels 32; Gaps 8;		
QY	10 RYLERYLLEA-KEANITTCGAHCSLNMENIV---PDKVNFVAMKRMV-VGOQAV-EV 63		
DB	118 KQIDLAIVSAVELNKLSPSCDSKAITYGVITISANADEKVALIAEAMEKAGNDGVITY 177		
QY	64 WQGLDAVE-AVLKQALLVNSQFW-----EPLQLHVDRAVSGLSRLTLL 109		
DB	178 EEGTGLONELEVYKMGQDFRGYLSFYINKPETGIVTELENPIYLVADKRISVREMLPIL 237		
QY	110 RALGAKREKISPPDASAPLPTITAD-TFRKLFRVYSNFLGKGLKLYTGEGCKRGDR 166		
DB	238 ESV-----AKSGRPILLISEDLGEGALATLVNMSMGIVKVAAYKAPGEGDR 284		
RESULT 14			
ID	CH60_BUCAP	STANDARD:	PRT: 551 AA.
AC	CH60_BUCAP	059177:	
DT	01-NOV-1997	(Rel. 35, Created)	

```

DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN      GROL OR MOP4 OR GROEL
OS      Buchnera aphidicola (subsp. Schizaphis graminum).
OC      Bacteria: Proteobacteria; gamma subdivision; Buchnera.
RX      NCBI_TaxId=98794;
RN      [1]
RA      APPARU H.K., Moriya S., Baumann P., Yoshikawa H., Ogasawara N.;
RT      "The nucleotide sequence of 60k, tdhf, groES and groEL genes of
RL      Buchnera aphidicola."
RL      Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RL      [2]
RF      SEQUENCE FROM N.A.
RA      MEDLINE=98184963; PubMed=9516544;
RA      Clark M.A., Baumann L., Baumann P.;
RT      "Sequence analysis of a 34.7-kb DNA segment from the genome of
RL      Buchnera aphidicola (endosymbiont of aphids) containing groEL, dnaA,
RL      the alp operon, gldA, and rno."
RL      Curr. Microbiol. 36:158-163(1998).
CC      -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC      PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC      CONDITIONS. (BY SIMILARITY).
CC      -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC      7 SUBUNITS (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; D85628; BAA12847.1; -
DR      EMBL; AF008210; AAC38095.1; -
DR      HSSP; P06139; 1JON.
DR      InterPro: IPR001844; Chaperlin_Cpn60.
DR      InterPro: IPR002423; Cpn60/TCF-1.
DR      Pfam: PF00118; cpn60_TCF1.1
DR      PRINTS: PR00298; CHAPERONIN60.
DR      PRINTS: PR00304; TCOMPLEXTCP1.
DR      PROSITE: PS00296; CHAPERONINS_CPN60; 1.
DR      Chaperone; ATP-binding.
SQ      SEQUENCE 551 AA; 58207 MW; 4C8176847ED04A5D CRC64;
Query Match
Best Local Similarity 9.5%; Score 81; DB 1; Length 551;
Matches 43; Conservative 33; Mismatches 70; Indels 32; Gaps 8;
OY      10 RYLEYELIEA-KEENITVTGCAEHCSLIENTV--PDTKVNFAWMKE-VGQAAV-EV 63
DB      121 RQIDAATSAAEELKNLSVPCSDSKAATQVGTISANADEKGALIAAMEKVGNDVTFV 180
OY      64 WQGLALISE-AVLGQALIVNSSQPV-----EPQLQHYKAVASGLSTLTIL 109
DB      181 EEGGLQLEELVYVGMQFDRCYLSPLPKPRTGIVLELNPYILMAKKISNVREMLPIL 240
OY      110 RALGAQKRAISPPDAASAPLRTTAP-TFKRLFRVSNFRLGKLTQYTGACRTGGR 166
DB      241 ESV-----AKSGKPLLTISEDLGEGALATLVVNMGRGIVKAAVAKAGFGDR 287
RESULT 15
POLG-HCVJ8
POLG-HCVJ8 STANDARD; PRT; 3033 AA.
AC      P26661;
DT      01-AUG-1992 (Rel. 23, Created)
DT      01-AUG-1992 (Rel. 23, Last annotation update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Genome polypeptide [contains: Capsid protein C (Core protein) (P22);

```

DE Envelope glycoprotein E1 (GP33) (GP33): Envelope glycoprotein E2
DE (GP68) (GP70) (NS1): Protein p7: Nonstructural protein NS2 (p21)
DE (EC 3.4.22.-): Protease/helicase NS3 (p70) (Hepatitisin)
DE (EC 3.4.21.98): Nonstructural protein NS4A (p4): Nonstructural protein
DE NS4B (p27): Nonstructural protein NS5A (p56): Nonstructural protein
DE NS5B (p66) (p70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
DE Hepatitis C virus (Isolate HCV-58) (HCV).
OC Viruses: ssRNA positive strand viruses, no DNA stage; Flaviviridae;
OC Hepatitis C virus
OX NCBI_TaxId=11115.
RN [1]
RP
RX MEDLINE=92230232; PubMed=1314459;
RA Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
RA Fukuda S., Tsuda F., Mishiro S.;
RT "Full-length sequence of a hepatitis C virus genome having poor
RT homology to reported isolates: comparative study of four distinct
RT genotypes.";
RL Virology 188:331-341(1992).
CC
CC -1 FUNCTION: "THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION."
CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC
CC -1 CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
CC precursor polyprotein, commonly with Asp or Glu in the p6
CC position, Cys or Thr in p1 and Ser or Ala in p1' or p1'' in the p6
CC
CC -1 CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA) (N).
CC
CC -1 SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPID-PROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA.
CC
CC -1 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: D10968; BAA01761.1; -
DR PIR: A40250; GNMV08.
DR HSSP: P27958; 1HEI.
DR MEROPS: S29.001; -
DR
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR002518; HCV_NS2.
DR InterPro: IPR004109; HCV_NS3.
DR InterPro: IPR000745; HCV_NS4A.
DR InterPro: IPR001490; HCV_NS4B.
DR InterPro: IPR002868; HCV_NS5A.
DR InterPro: IPR002166; HCV_RDRP.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR Pfam: PF00998; HCV_RDRP; 1.
DR Pfam: PF01001; HCV_NS4B; 1.
DR Pfam: PF01006; HCV_NS4A; 1.
DR Pfam: PF01506; HCV_NS5A; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR ProDom: PD186062; HCV_NS1; 1.
DR PolyProtein: Glycoprotein; Transferase; RNA-directed RNA polymerase;
KW Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
KW Transmembrane; Nonstructural
INT1_MET 1
FT REMOVED FROM CAPSID PROTEIN C BY THE
CELLULAR AMINOPEPTIDASE.

```
FT CHAIN 1 115 CAPSID PROTEIN C (POTENTIAL).
FT CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).
FT CHAIN 192 383 MAJOR ENVELOPE PROTEIN E (POTENTIAL).
FT CHAIN 384 733 NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
FT CHAIN 734 1010 NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
FT CHAIN 1011 1619 PROTEASE/HELICASE NS3 (POTENTIAL).
FT CHAIN 1620 1866 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
FT CHAIN 1867 2017 NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
FT CHAIN 2018 3033 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
FT TRANSMEM 347 369 POTENTIAL.
FT ACT_SITE 1087 1087 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 1111 1111 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 1169 1169 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT NP_BIND 1234 1241 ATP (POTENTIAL).
FT SITE 1320 1323 DECH BOX.
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 477 477 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 558 558 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 578 578 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 649 649 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1091 1091 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2038 2038 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2359 2359 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2811 2811 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 3033 AA; 330177 MW; 1A173E7E381FD1A CRC64;
```

Query Match 9.5%; Score 80.5; DB 1; Length 3033;
Best Local Similarity 26.6%; Pred. No. 24;

Matches 41; Conservative 22; Mismatches 64; Indels 27; Gaps 7;

```
QY 5 LICDSRFLERYLLEAKENITFGCAE--HCSINENITV-PDTKVNIFYAMKRM----- 54
DB 1659 IMTSSWVLGAGVLAANAICLATGICISITGRHLNDRVYVAPDKELIYAFDEMECCASK 1718
QY 55 ----EYGOQAVEVMQG--LALLSEAVLRGQALLVNSSQPEPLQ---LHYDKAVSGIIRS 104
DB 1719 AALIEGQRMAEMLKSKIIGLLQOATROADIQPAIQSSMPKLECFWAKHMMNFISGIQY 1778
QY 105 LTTLLRALG----AQKEAISPPDAASAAPLRTIT 134
DB 1779 LAGLSTLPGNPVAVSMMAFS---AALTSPILPTST 1809
```

Search completed: January 2, 2003, 15:11:53
Job time : 10.5257 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:09:40 ; Search time 26.5801 Seconds
(without alignments)
1286.823 Million cell updates/sec

Title: US-09-853-731-2

Perfect score: 851
Sequence: 1 APRRLICDSRVLERYLLEAK.....NFRGKRLKLYTEACRTGDR 166

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP-archaea:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP-invertebrate:*
6: SP-mammal:*
7: SP-mhc:*
8: SP-organella:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-virus:*
13: SP-vertebrate:*
14: SP-unclassified:*
15: SP-virus:*
16: SP-bacteriap:*
17: SP-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	685.5	80.6	195	09GKA3	09gka3 oryctolagus
2	685.5	80.6	195	09GKA2	09gka2 oryctolagus
3	683	80.3	194	09MYM8	09mym8 sus scrofa
4	188	22.1	50	09QV40	09qv40 rattus sp.
5	88	10.3	323	16 082DC8	082dc8 salmoneilla
6	87.5	10.3	346	16 082KZ4	082kz4 salmoneilla
7	87.5	10.3	346	16 082ZM5	082zm5 salmoneilla
8	85	10.0	339	16 09H2M7	09h2m7 pseudomonas
9	85	10.0	3722	2 P94873	P94873 lysobacter
10	83.5	9.8	296	16 082AY4	082ay4 yersinia pe
11	83	9.8	548	2 093T48	093t48 buchiera ap
12	83	9.8	548	2 093N35	093n35 buchiera ap
13	82.5	9.7	3033	12 093DH6	093dh6 hepatitis c
14	82	9.6	815	10 09FK91	09fk91 arabidopsis
15	80	9.4	548	2 093N34	093n34 buchiera ap
16	79.5	9.3	389	2 09AN18	09an18 bradyrhizob

17	79.5	9.3	1829	16 086808	086808 streptomyce
18	79.5	9.3	3033	12 09JZAI	09jzai hepatitis c
19	79	9.3	480	10 043380	043380 avena sativ
20	79	9.3	2364	5 Q22896	Q22896 caenorhabd
21	78.5	9.2	813	16 082989	082989 salmoneilla
22	78.5	9.2	3019	12 092529	092529 hepatitis c
23	78	9.2	348	16 086751	086751 streptomyce
24	78	9.2	378	17 082ZK3	082zk3 pyrobaculum
25	78	9.2	455	16 08UB09	08ub09 agrobacteri
26	78	9.2	548	2 093K25	093k25 buchiera ap
27	77.5	9.1	242	16 08Y76	08y76 anabaena sp
28	77.5	9.1	310	16 09L033	09l033 streptomyce
29	77.5	9.1	383	16 0920V7	0920v7 rhizobium m
30	77.5	9.1	451	16 08XSE8	08xse8 ralsstonia s
31	77.5	9.1	1003	10 09FIR4	09fir4 arabidopsis
32	77.5	9.1	2536	11 091XW2	091xw2 mus muscul
33	77	9.0	379	2 09L9P5	09l9p5 streptomyce
34	77	9.0	447	2 P72270	P72270 rhodococcus
35	77	9.0	451	16 P74054	P74054 synechocyst
36	76.5	9.0	154	16 09RGA7	09rga7 xyella fas
37	76.5	9.0	425	16 08YF20	08yf20 bruceella me
38	76.5	9.0	619	3 096VB7	096vb7 botrytis cl
39	76.5	9.0	871	5 09NB04	09nb04 drosophila
40	76.5	9.0	1315	3 08X215	08x215 botrytis cl
41	76.5	9.0	1315	3 08X1E7	08x1e7 botrytis cl
42	75.5	8.9	486	2 09F2G8	09f2g8 streptococ
43	75.5	8.9	637	16 055413	055413 synechocyst
44	75	8.8	308	10 094I41	094i41 oryza sativ
45	75	8.8	397	8 09TNG5	09tng5 coriaria te

ALIGNMENTS

RESULT 1
ID 09GKA3 PRELIMINARY; PRT; 195 AA.
AC 09GKA3;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Erythropoietin.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;

OX [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-21290682; PubMed-11396976;

RA Vialita A., Wu D., Margalith M., Hobart P.;

RT "Rabbit EPO Gene and cDNA: Expression of Rabbit EPO after

RT Intramuscular Injection of pDNA.";

RL Biochem. Biophys. Res. Commun. 284:823-827(2001).

DR EMBL; AF290943; AAC36961.1; -

DR HSSP; P01588; ICN4.

DR InterPro; IPR001323; EPO.TPO.

DR InterPro; IPR003013; Erythropo.

DR Pfam; PF00758; EPO.TPO.1.

DR PRINTS; PR00272; ERYTHROPTN.

DR PROSITE; PS00817; EPO.TPO.1.

SO SEQUENCE 195 AA; 21053 MW; 0999DAYD852713F3 CRC64;

Query Match 80.6%; Score 685.5; DB 6; Length 195;
Best Local Similarity 81.4%; Pred. No. 8.8e-61;

Matches 136; Conservative 12; Mismatches 18; Indels 1; Gaps 1;

QY	1	APRRLICDSRVLERYLLEAKENITTCAGHCSLNTITVPTDKVFEYMKRMEVGOQA	60
DB	29	APARLICDSRVLERYLLEAKENITTCAGHCSLNTITVPTDKVFEYMKRMEVGOQA	88
QY	61	VEVWOGIALISEAVLRGQALLVNSQPEWPLQLHVDKAVSGLSLTLRALGAKQKAIS	120
DB	89	VEVWOGIALISEAVLRGQALLVNSQPEWPLQLHVDKAVSGLSLTLRALGAKQKAIS	148

OY 121 PPDA-SAAPLRTTADTFKRLFRVYSNPLRGKLTLYTGACRTRDR 166
 DB 149 PPEAASSAAPLRTVAADTLCKLFRVYSNPLRGKLTLYTGACRTRDR 195

RESULT 2

O9GKA2 PRELIMINARY; PRT; 195 AA.
 AC O9GKA2;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Erythropoietin.
 OS Erythrocyte cuniculus (Rabbit).
 OC Mammalia; Euteria; Lagomorpha; Leporidae; Oryctolagus.
 NCBI_Taxid=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21290682; PubMed=11396976;
 RA Viala A., Ma D., Margalich M., Hobart P.;
 RT Rabbit EPO gene and cDNA: Expression of Rabbit EPO after
 RL Intramuscular Injection of pDNA.
 RL Biochem. Biophys. Res. Commun. 284:823-827(2001).
 DR HSP, P01588; ICN4.
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003013; Erythropo.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PRINTS: PR00272; ERYTHROPTN.
 DR PROSITE: PS00817; EPO_TPO; 1.
 SQ SEQUENCE 195 AA; 21025 MW; 1F1DC7F403A03BC CRC64;

Query Match 80.6%; Score 685.5; DB 6; Length 195;
 Best Local Similarity 81.4%; Pred. No. 8.6e-61;
 Matches 136; Conservative 12; Mismatches 18; Indels 1; Gaps 1;

OY 1 APRLLICDSVRLERYLLEAKENITTCGAHCNSLNNITVPDTRVNFYAMKRMVGOQA 60
 DB 29 APRLLICDSVRLERYLLEAKENITTCGAHCNSLNNITVPDTRVNFYAMKRMVGOQA 88
 OY 61 VEWQGLALISEAVLRGQALLVNSQSPWEPLOLHVDAVSGRLSTTLRALGAQKEAIS 120
 DB 89 VEWQGLALISEAVLRGQALLVNSQSPWEPLOLHVDAVSGRLSTTLRALGAQKEAIS 148
 OY 121 PPDA-SAAPLRTTADTFKRLFRVYSNPLRGKLTLYTGACRTRDR 166
 DB 149 PPEAASSAAPLRTVAADTLCKLFRVYSNPLRGKLTLYTGACRTRDR 195

RESULT 3

O9MYM8 PRELIMINARY; PRT; 194 AA.
 AC O9MYM8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Erythropoietin precursor.
 GN EPO.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_Taxid=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NORWEGIAN LANDRACE; TISSUE=KIDNEY;
 RA David B., Harbitz I.;
 RT "The porcine erythropoietin gene: cDNA and genomic sequences and
 expression analyses."
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ249745; CAB96416.1; -;
 DR EMBL; AJ249746; CAB96417.1; -;

DR HSP, P01588; ICN4.
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003013; Erythropo.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PRINTS: PR00272; ERYTHROPTN.
 DR PROSITE: PS00817; EPO_TPO; 1.
 KW Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 194
 SQ SEQUENCE 194 AA; 21303 MW; 77881A6F620E1C CRC64;

Query Match 80.3%; Score 683; DB 6; Length 194;
 Best Local Similarity 82.1%; Pred. No. 1.6e-60;
 Matches 138; Conservative 7; Mismatches 21; Indels 2; Gaps 1;

OY 1 APRLLICDSVRLERYLLEAKENITTCGAHCNSLNNITVPDTRVNFYAMKRMVGOQA 60
 DB 27 APRLLICDSVRLERYLLEAKENITTCGAHCNSLNNITVPDTRVNFYAMKRMVGOQA 86
 OY 61 VEWQGLALISEAVLRGQALLVNSQSPWEPLOLHVDAVSGRLSTTLRALGAQKEAIS 120
 DB 87 VEWQGLALISEAVLRGQALLVNSQSPWEPLOLHVDAVSGRLSTTLRALGAQKEAIS 146
 OY 121 PPDA-SAAPLRTTADTFKRLFRVYSNPLRGKLTLYTGACRTRDR 166
 DB 147 LPDAPSSAPLRTVAADTLCKLFRVYSNPLRGKLTLYTGACRTRDR 194

RESULT 4

O9QV40 PRELIMINARY; PRT; 50 AA.
 AC O9QV40;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Erythropoietin (Fragment).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=94115047; PubMed=7764337;
 RA Okano M., Suga H., Masuda S., Nagao M., Narita H., Ikura K.,
 RL Sasaki K.;
 DR Biosci. Biotechnol. Biochem. 57:1882-1885(1993).
 DR HSP, P01588; IER.
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003013; Erythropo.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PRINTS: PR00272; ERYTHROPTN.
 SQ SEQUENCE 50 AA; 5587 MW; 70B44A8BFE016034 CRC64;

Query Match 22.1%; Score 188; DB 11; Length 50;
 Best Local Similarity 78.0%; Pred. No. 1.2e-11;
 Matches 39; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 APRLLICDSVRLERYLLEAKENITTCGAHCNSLNNITVPDTRVNFYAMKRMVGOQA 60
 DB 1 APRLLICDSVRLERYLLEAKENITTCGAHCNSLNNITVPDTRVNFYAMKRMVGOQA 86

RESULT 5

O8ZDC8 PRELIMINARY; PRT; 323 AA.
 AC O8ZDC8;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Ribonucleoside-diphosphate reductase 2 beta chain (EC 1.17.4.1).
 GN NRDF OR YPO2648.
 OS Versinia pestis.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

RC STRAIN=NORWEGIAN LANDRACE; TISSUE=KIDNEY;
 RA David B., Harbitz I.;
 RT "The porcine erythropoietin gene: cDNA and genomic sequences and
 expression analyses."
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ249745; CAB96416.1; -;
 DR EMBL; AJ249746; CAB96417.1; -;


```

GN UDP-N-acetylpyruvoylglycosamine reductase.
OS MURB OR PA2977.
OC Pseudomonas aeruginosa.
OC Pseudomonas.
OC Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC NCBI_TaxID=287;
RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.C., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M., Glickman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garbey R.J., Goulet L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Kas A., Larbig K., Lim R.M.,
RA Reizer J., Salas M.H., Hancock R.E.W., Lory S., Olson M.V.,
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000);
DR EMBL: AF004723; AAC06305.1; -.
DR HSSP: P08373; 2MBR.
DR Interpro: IPR003170; MuB.
DR Interpro: IPR001575; Oxid_FAD_bind.
DR Pfam: PF01565; FAD_binding_4; 1.
DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
DR Complete proteome.
KM SEQUENCE 339 AA; 37627 MW; C0C8E9F2938FE27 CRC64;

Query Match 10.0%; Score 85; DB 16; Length 339;
Best Local Similarity 23.0%; Pred. No. 2.9;
Matches 42; Conservative 22; Mismatches 49; Indels 70; Gaps 8;

QY 31 EHCSLNE-NITVPPTKYNFYAKRMFYGGQAVFYWGGLALISPAVIRGQALLV----- 82
DB 7 EHCSLKYNTFEGIDVARRLAHARDE----ADVREALAHAR--RGJPLVLITGGGSNL 58
QY 83 -----NSSQPEP-IQLHYDKAVSGLSRTITLL 109
DB 59 LITRDVEALVLRMASGGRIVSDAADSIVLEAAGAMPFVOWSLERLAGLGNLSLT- 117
QY 110 RALGAQKEAISPPDASAPLRTITA-----DTEKLFRRVSNFLRKLYTGBACT 163
DB 118 -----PGTVGAPPMONIGAVGLMDVDLSIAL--DQDGLRFRDQACRF 163
QY 164 GDR 166
DB 164 GVR 166

RESULT 9
P94873 PRELIMINARY; PRT; 3722 AA.
AC P94873;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Alpha-aminoadipyl-cysteinyl-valine synthetase.
GN PCBAAB.
OS Lysobacter lactamgenus.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Lysobacter.
OX NCBI_TaxID=39596;
RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YK90;
RX MEDLINE=96275949; PubMed=8737573;
RA Kimura H., Miyashita H., Sumino Y.;
RA Kinugata and expression in Pseudomonas putida of the gene cluster
RT involved in cephalosporin biosynthesis from lysobacter lactamgenus
RT YK90.
RL Appl. Microbiol. Biotechnol. 45:490-501(1996).

```

```

DR EMBL: D50308; BAA0846.1; -.
DR HSSP: P14687; 1AND.
DR Interpro: IPR002490; AACRNA_LigaseII.
DR Interpro: IPR002499; AMP-bind.
DR Interpro: IPR001242; AMP-bind.
DR Interpro: IPR003972; DNase.
DR Interpro: IPR003860; Spentate-attach.
DR Interpro: IPR003979; Tel-esterase.
DR Interpro: IPR010311; Thiocetate.
DR Pfam: PF00501; AMP-binding_3.
DR Pfam: PF00668; Condensation_3.
DR Pfam: PF00550; pp-binding_3.
DR Pfam: PF00975; Thiocetate_3.
DR PROSITE: PS00179; AA-TRNA-LIGASE_IL_1; UNKNOWN_1.
DR PROSITE: PS50075; ACE_DOMAIN_3.
DR PROSITE: PS00455; AMP_BINDING_1.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_2.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_2.
DR Phosphopantetheine.
KM SEQUENCE 3722 AA; 411607 MW; 3597B3483463809B CRC64;

Query Match 10.0%; Score 85; DB 2; Length 3722;
Best Local Similarity 22.1%; Pred. No. 61;
Matches 34; Conservative 18; Mismatches 50; Indels 52; Gaps 5;

QY 16 ILEAKEANTITGCAHCSINENTIVPTKYNFYAKRMFYGGQAVFY--WGGLALLS 71
DB 2138 ILEGATAGTLP-----PIPATYADFSYWKQSDQRDALDFDYO----- 2178
QY 72 EAVLRGQALLVNSSQPEP-IQLHYDKAVSGLSRTITLLRALGAQ-----KEAISPPDAS 126
DB 2179 -----RSLAGWQPIQLPLD-----HARPAFDYLGRIEIVFDATT 2214
QY 127 AAPLRTTADPFRKLFRRVSNFLRKLYTGEA 160
DB 2215 CDQLRVLTAQTRTSFFSVLAAVYLLTKAVSGS 2248

RESULT 10
Q8ZAV4 PRELIMINARY; PRT; 296 AA.
AC Q8ZAV4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative 2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60).
GN YPO3648.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=11586360;
RA Parrhill J., Wren B.W., Thomson N.R., Tibball R.W., Holden M.T.G.,
RA Prentice M.B., Sepahnia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Fellwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
RT Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001);
DR EMBL: A014158; CAC93117.1; -.
DR Interpro: IPR002204; 3hydroxyisobut_dh.
DR Interpro: IPR001744; 6eGD.
DR Pfam: PF03446; NAD_binding_2; 1.
DR PROSITE: PS00895; 3-HYDROXYISOBUT_DH; UNKNOWN_1.
DR Oxidoreductase; Hypothetical protein; Complete proteome.
KM SEQUENCE 296 AA; 32140 MW; C883AFCC686429D CRC64;

Query Match 9.8%; Score 83.5; DB 16; Length 296;

```



```

DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00998; HCV NS5b; 1.
DR Pfam: PF0186062; HCV NS1; 1.
DR SMART: SM00487; DEXDC; 1.
DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN_1.
DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; RNA directed RNA polymerase; Transmembrane.
FT CHAIN 1
FT CHAIN 192
FT CHAIN 383
FT CHAIN 750
FT CHAIN 751
FT CHAIN 813
FT CHAIN 814
FT CHAIN 1030
FT CHAIN 1031
FT CHAIN 1061
FT CHAIN 1062
FT CHAIN 1115
FT CHAIN 1176
FT CHAIN 1242
FT CHAIN 1243
FT CHAIN 3033
SQ SEQUENCE 3033 AA; 329981 MW; 68183FED09087284 CRC64;

Query Match
Best Local Similarity 27.0%; Score 82.5; DB 12; Length 3033;
Matches 40; Conservative 21; Mismatches 60; Indels 27; Gaps 7;

QY 11 VLERYLEAKAEENITTCGAE--HCSLNENTIV-PDTKVFYAMKRM-----EV 56
Db 1665 VLAGGVLAAYVAYGLATGCGSTIGRIHLNDQVYVAPREKILYALFEMECCASKALIEE 1724

QY 57 GQOAEVWQG--LALSEAVIRGOALLVNSQSPWEPLQ---LHVDKAVSGLSLTLLR 110
Db 1725 GQRAEMKIKSLTGLLQDAARKQADIDPAWQSSMPKIEQFARHWNFTISGLIYLAGLST 1784

QY 111 ALG----AQKRAISPPDASAPLRTTT 134
Db 1785 LPGNPAVASMARS--AALTSLPLST 1809

RESULT 14
ID 09FR91 PRELIMINARY; PRT; 815 AA.
AC 09FR91;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 19, Last sequence update)
DE Seed maturation protein PM38 protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 111
RF SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RK MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Nabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Physically assigned P1 and TAC clones."
RL DNA Res. 5:203-216(1998).
DR EMBL: AB012244; BAB09119.1;
DR InterPro: IPR002086; Aldehyde_dehydr.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR001290; PARP.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF00644; PARP; 1.
DR SMART: SM00292; BRCT; 1.
DR PROSITE: PS00170; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE: PS00172; BRCT; 1.
SQ SEQUENCE 815 AA; 91534 MW; 4B1B602057D46B60 CRC64;

```

```

Query Match
Best Local Similarity 9.6%; Score 82; DB 10; Length 815;
Matches 42; Conservative 29; Mismatches 52; Indels 70; Gaps 12;

QY 21 EAENITTCGA-----EHCSLNENTIVPDKVFYAMKRMVEG----QOAEVW 64
Db 328 EKDGLIYNCAFSCIDLGKGRNEYC-IMQLVTPDSLNMT-FRKGVDGDPNAEERLEW 385

QY 65 OGALISEAVIRGOALLV-----NSQSPWE-----PLQ---LHVDKAVSGL 102
Db 386 ED---EFAAIKFALEFETIAGNEFEPEWEREKIQKKRPFETMDGDIEVSGALGL 441

QY 103 RSL-----TTLRALGAQ-----KEAISPDASAPLRTITADTFEKL 141
Db 442 RQLGTSAMCKLDSFVAFNFTKVLGQGEIYNVALMELGIDPD---LPMGLMTDTHKRC 497

QY 142 FRYYSNPLGKIK 154
Db 498 EYVLEFVE-KVK 509

RESULT 15
ID 093N34 PRELIMINARY; PRT; 548 AA.
AC 093N34;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE groEL-related molecular chaperonin Syml.
GN Syml.
OS Buchnera aphidicola (subsp. Rhopalosiphum padi).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=98793;
RN 111
RF SEQUENCE FROM N.A.
RC "Rhopalosiphum padi endosymbiont GroEL-related molecular chaperonin
RT Syml (syml) gene."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF387864; AAK62971.1;
DR InterPro: IPR001844; Chaperonin_Cpn60.
DR InterPro: IPR002423; Cpn60_TCP-1.
DR Pfam: PF00118; Cpn60_TCP1; 1.
DR PROSITE: PS00296; CHAPERONINS_CPN60; UNKNOWN_1.
KW ATP-binding; Chaperone.
SQ SEQUENCE 548 AA; 57752 MW; C0296E8B6EF2794 CRC64;

Query Match
Best Local Similarity 9.4%; Score 80; DB 2; Length 548;
Matches 43; Conservative 32; Mismatches 71; Indels 32; Gaps 8;

QY 10 RVLERYLEA-KEAENITTCGAHCSLNENTIV--PDTKVFYAMKRM-VGQOAV-EV 63
Db 118 RGLDKAVISAVELKMLSPCSDSKAITOVGTISANADKVCALIAEMERKNGDVGITV 177

QY 64 WGLALISE-AVLRGOALVNSQSPW-----EPLQVHVDKAVSGLSLTLL 109
Db 178 EBSGTGLNELEVVKQCFDRLGYLSFYINKPEPTGVLEENPYITLADKISNVREMLPIL 237

QY 110 RALGAQKEAISPDASAPLRTITAD-TFRKLFYVSNFSLGKIKILYTGACRTGDR 166
Db 238 ESV-----AKSGKPLITISEDLEGEALATLVNSTRGIYKVAAYKAPFGDR 284

Search completed: January 2, 2003, 15:12:56
Job time: 28.5801 secs

```